Discussion Paper Series

No.197

Asymptotic Properties of the Efficient Estimators for Cointegrating Regression Models with Serially Dependent Errors

Eiji Kurozumi
Kazuhiko Hayakawa

December 2006

Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences
A 21st-Century COE Program

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan
http://hi-stat.ier.hit-u.ac.jp/

Asymptotic Properties of the Efficient Estimators for Cointegrating Regression Models with Serially Dependent Errors

EIJI KUROZUMI[1,2] KAZUHIKO HAYAKAWA[3]

Department of Economics

Hitotsubashi University

December, 2006

Abstract

In this paper, we analytically investigate three efficient estimators for cointegrating regression models: Phillips and Hansen's (1990) fully modified OLS estimator, Park's (1992) canonical cointegrating regression estimator, and Saikkonen's (1991) dynamic OLS estimator. First, by the Monte Carlo simulations, we demonstrate that these efficient methods do not work well when the regression errors are strongly serially correlated. In order to explain this result, we assume that the regression errors are generated from a nearly integrated autoregressive (AR) process with the AR coefficient approaching 1 at a rate of $1/T$, where T is the sample size. We derive the limiting distributions of the three efficient estimators as well as the OLS estimator and show that they have the same limiting distribution under this assumption. This implies that the three efficient methods no longer work well when the regression errors are strongly serially correlated. Further, we consider the case where the AR coefficient in the regression errors approaches 1 at a rate slower than $1/T$. In this case, the limiting distributions of the efficient estimators depend on the approaching rate. If the rate is slow enough, the efficiency is established for the three estimators; however, if the approaching rate is relatively fast, they have the same limiting distribution as the OLS estimator. This result explains why the effect of the efficient methods diminishes as the serial correlation in the regression errors gets stronger.

JEL classification: C13; C22

Key Words: Cointegration; second-order bias; fully modified regressions; canonical cointegrating regressions; dynamic ordinary least squares regressions

---

[1]Kurozumi's research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology under Grants-in-Aid No. 17203016 and 18730142.

[2]Correspondence: Eiji Kurozumi, Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, 186-8601, Japan. E-mail: kurozumi@stat.hit-u.ac.jp

[3]Hayakawa is a JSPS research fellow and acknowledges its financial support.

1. Introduction

Since the seminal work of Engle and Granger (1987), cointegrating regressions have become one of the standard tools in analyzing integrated variables. With regard to the estimation of cointegrating regression models, it is well known that the ordinary least squares (OLS) estimator contains the second-order bias, comprising the endogeneity bias and the non-centrality bias, when the I(1) regressors are endogenous and/or the regression errors are serially correlated. Thus, several efficient methods for the estimation of the cointegrating regressions have been proposed in the literature. Phillips and Hansen (1990) proposed a nonparametric correction for the OLS estimator; their method is known as the fully modified regression (FMR) method, which has been further developed by Phillips (1995) and Kitamura and Phillips (1997). Park (1992) proposed the canonical cointegrating regression (CCR) method, which is also based on a nonparametric correction similar to the FMR method. However, the CCR method eliminates the non-centrality bias in a different manner. On the other hand, Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993) considered a parametric correction by adding leads and lags of the first difference of the I(1) variables as regressors; this method is known as the dynamic ordinary least squares (DOLS) method. These three efficient estimators—the FMR, CCR, and DOLS estimators—are asymptotically equivalent, and as proved by Saikkonen (1991), they are efficient.

However, the finite sample behavior of these estimators is fairly different as reported by, for example, Inder (1993), Montalvo (1995), Cappuccio and Lubian (2001), and Christou and Pittis (2002) through Monte Carlo simulations. The first two papers recommend using the DOLS type approach to eliminate the second-order bias of the OLS estimator, while the last paper demonstrated that the FMR estimator outperforms the DOLS estimator in view of the bias; the answer to the question of which estimator performs best in finite samples is inconclusive. It seems that the performance of the three efficient estimators is fairly dependent on the data generating process used in Monte Carlo simulations, as pointed out by Cappuccio and Lubian (2001). However, these Monte Carlo simulations commonly suggest that the efficient estimation methods break down and perform very poorly when the cointegrating regression errors are strongly serially correlated. Although the finite sample performance of the FMR and CCR estimators may improve if the prewhitening method by Andrews and Monahan (1992), which has been further modified by Sul, Phillips, and Choi (2005), is used to estimate the long-run variance, a large bias still remains in the estimator as shown in the later section.

In this paper, we analytically explain the poor performance of the three efficient estimators with a strong serial correlation. We first consider the first-order autoregressive (AR) process for the regression errors and assume the local-to-unity system in which the AR coefficient approaches one at the rate of $1/T$, where T is the sample size. This system has been considered in the literature in order to investigate the local asymptotic behavior of unit root and cointegration tests. We will show that the FMR, CCR, and DOLS estimators have the same limiting distribution as the OLS estimator under this system. This implies that the three efficient methods break down and the three estimators perform as poorly as the OLS estimator. Next, we introduce the local-to-unity system in which the AR coefficient approaches 1 at a rate slower than $1/T$. This type of local-to-unity system is also considered by Phillips and Magdalinos (2005, 2007) and Giraitis and Phillips (2006). We will show that the limiting distributions of the efficient estimators change depending on the approaching speed of the AR coefficient. Intuitively, the three efficient methods can eliminate the second-order bias of the OLS estimator if the AR coefficient approaches 1 slowly enough, while these methods no longer work when the approaching speed is very fast. For the intermediate case, the second order bias of the OLS estimator is partially eliminated by the efficient methods; however, a part of the bias still remains. This result explains why the effect of the efficient methods diminishes as the serial correlation in the regression errors becomes stronger. We will show that the result depends on a relation between the approaching speed of the AR coefficient and the diverging rate of the bandwidth parameter used for the estimation of the long-run variance in the FMR and CCR methods or the diverging rate of the lead-lag truncation parameter used in the DOLS method.

The remainder of this paper is organized as follows. In Section 2, we briefly review the

FMR, CCR, and DOLS methods, and in Section 3 we provide the motivating simulation results. In Section 4, we investigate the asymptotic property of the three efficient estimators as well as the OLS estimator under different types of local-to-unity systems. Concluding remarks are provided in Section 5.

2. Review of the Efficient Estimation Methods

This section reviews the three efficient estimators for cointegration regression models. Let us consider the following model:

$$\begin{aligned} y_t &= \mu + \beta' x_t + u_{1t} &= \theta' z_t + u_{1t} \\ \Delta x_t &= u_{2t} \end{aligned} \tag{1}$$

for $T = 1, \cdots, T$, where $\theta = (\mu, \beta')'$, $z_t = (1, x_t')'$, and y_t and x_t are observed time series with 1 and n dimensions, respectively. For $u_t = [u_{1t}, u_{2t}']'$, we assume that the functional central limit theorem (FCLT) can be applied as follows:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} u_t \Rightarrow W(r) = \begin{bmatrix} W_1(r) \\ W_2(r) \end{bmatrix}$$

for $0 \leq r \leq 1$, where $W(\cdot)$ is a Brownian motion on $[0, 1]$ with a variance-covariance matrix Ω ($W(\cdot) \sim BM(\Omega)$) and $\Rightarrow$ signifies weak convergence of the associated probability measures. Note that the long-run variance of u_t and its one-sided version can be expressed as

$$\Omega = \Sigma_u + \Pi + \Pi' \quad \text{and} \quad \Lambda = \Sigma_u + \Pi,$$

$$\text{where} \quad \Sigma_u = \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} E(u_t u_t') \quad \text{and} \quad \Pi = \lim_{T \to \infty} T^{-1} \sum_{j=1}^{T-1} \sum_{t=1}^{T-j} E(u_t u_{t+j}').$$

We partition Ω and Λ conformably with u_t as

$$\Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \Omega_{22} \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \Lambda_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \Lambda_2 \end{bmatrix}. \tag{2}$$

It is known that the OLS estimator of θ, denoted by $\hat{\theta}$, is consistent but inefficient in general. The centered OLS estimator with a normalizing matrix $D_T = \text{diag}\{\sqrt{T}, TI_n\}$ weakly converges to

$$D_T(\hat{\theta} - \theta) \Rightarrow \left(\int_0^1 W_2(r) W_2'(r) dr \right)^{-1} \left(\int_0^1 W_2(r) dW_1(r) + \lambda_{21} \right) \tag{3}$$

and we can observe that this limiting distribution contains the second-order bias from the correlation between $W_1(\cdot)$ and $W_2(\cdot)$ and the non-centrality parameter λ_{21}. As explained in Phillips and Hansen (1990) and Phillips (1995), the former bias arises from the endogeneity of the I(1) regressor x_t while the non-centrality bias comes from the fact that the regression errors are serially correlated.

In order to eliminate the second-order bias, Phillips and Hansen (1990) proposes the FMR estimator, which is defined as

$$\hat{\theta}_{FMR} = \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \left(\sum_{t=1}^{T} z_t y_t^{+} - T\hat{J}^{+} \right), \tag{4}$$

$$\text{where} \quad y_t^{+} = y_t - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1} u_{2t} \quad \text{and} \quad \hat{J}^{+} = \begin{bmatrix} 0 \\ \hat{\lambda}_{21} - \hat{\Lambda}_{22}\hat{\Omega}_{22}^{-1}\hat{\omega}_{21} \end{bmatrix},$$

with $\hat{\omega}_{12}$, $\hat{\Omega}_{22}$, $\hat{\lambda}_{21}$, and $\hat{\Lambda}_{22}$ being consistent estimators of ω_{12}, Ω_{22}, λ_{21}, and Λ_{22}, respectively. It can be shown that the correction term for y_t is concerned with the correction for the endogeneity bias while $\hat{J}^{+}$ eliminates the non-centrality bias.

In order to define the CCR estimator, we first modify y_t and x_t such that

$$y_t^{*} = y_t - \left(\hat{\beta}'\hat{\Lambda}_2\hat{\Sigma}_u^{-1} + [0, \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}] \right) \hat{u}_t \quad \text{and} \quad z_t^{*} = (1, x_t^{*'})' \quad \text{with} \quad x_t^{*} = x_t - \hat{\Lambda}_2\hat{\Sigma}_u^{-1}\hat{u}_t,$$

where $\hat{\beta}$ is the OLS estimator of β and $\hat{u}_t = [\hat{u}_{1t}, \Delta x_t']'$ consists of the OLS residuals and the first difference of the I(1) regressors. Then, the CCR estimator proposed by Park (1992) is defined as

$$\hat{\theta}_{CCR} = \left(\sum_{t=1}^{T} z_t^{*} z_t^{*'} \right)^{-1} \left(\sum_{t=1}^{T} z_t^{*} y_t^{*} \right). \tag{5}$$

The CCR method uses the same principle as the FMR method in order to eliminate the endogeneity bias, while it deals with the non-centrality parameter in a different manner.

Contrary to the non-parametric approaches taken by the FMR and CCR methods, the DOLS method is based on parametric regressions. Phillips and Loretan (1991), Saikkonen (1991), and Stock and Watson (1993) propose to augment the leads and lags of the first difference of x_t as regressors and to estimate

$$y_t = \theta' z_t + \sum_{j=-K}^{K} \pi_j' \Delta x_{t-j} + \dot{u}_{1t}. \tag{6}$$

The DOLS estimator is defined as the OLS estimator of θ for (6):

$$\hat{\theta}_{DOLS} = \left(\sum_{t=K+1}^{T-K} \tilde{z}_t \tilde{z}_t' \right)^{-1} \left(\sum_{t=K+1}^{T-K} \tilde{z}_t \tilde{y}_t \right), \tag{7}$$

where $\tilde{z}_t$ and $\tilde{y}_t$ are regression residuals of z_t and y_t on $w_t = (u_{2,t+K}', ..., u_{2,t-K}')'$, respectively. The regression form (6) is based on the fact that under some regularity conditions, the regression errors u_{1t} in (1) can be expressed as

$$u_{1t} = \sum_{j=-\infty}^{\infty} \pi_j' u_{2t-j} + v_t, \tag{8}$$

where $\sum_{j=-\infty}^{\infty} ||\pi_j|| < \infty$, with $|| \cdot ||$ being the standard Euclidian norm; further, v_t is uncorrelated with u_{2t-j} for all j. For details, see Brillinger (1981). From (8), we observe that

$$\dot{u}_{1t} = v_t + \sum_{|j|>K} \pi_j' u_{2t-j}. \tag{9}$$

The uncorrelatedness of v_t with all the leads and lags of u_{2t} is an important property to prove that the DOLS method successfully eliminates the second-order bias of the OLS Estimator.

As explained in Phillips and Hansen (1990), Saikkonen (1991), and Park (1992), these three efficient estimators have an identical limiting distribution that is given by

$$D(\hat{\theta}_E - \theta) \Rightarrow \left(\int_0^1 W_2(r) W_2'(r) dr \right)^{-1} \int_0^1 W_2(r) dW_{1\cdot 2}(r), \tag{10}$$

where $\hat{\theta}_E = \hat{\theta}_{FMR}$, $\hat{\theta}_{CCR}$, and $\hat{\theta}_{DOLS}$ and $W_{1\cdot 2}(\cdot) \sim BM(\omega_{1\cdot 2})$ is independent of $W_2(\cdot)$ with $\omega_{1\cdot 2} = \omega_{11} - \omega_{12}\Omega_{22}^{-1}\omega_{21}$. Then, we observe that the three efficient methods can eliminate both the endogenous bias and the non-centrality parameter. Moreover, Saikkonen (1991) showed that this limiting distribution is efficient in a certain class of estimators.

3. Finite Sample Evidence

This section investigates the finite sample performance of the three efficient estimators as well as the OLS estimator through Monte Carlo simulations. In the simulations, we focus on the effect of the serial correlation in the cointegrating regression errors, and thereby, we consider the following simple data generating process:

$$y_t = \mu + \beta x_t + u_{1t}, \quad x_t = x_{t-1} + u_{2t},$$

where x_t is a scalar unit root process. The error term $u_t = [u_{1t}, u_{2t}]'$ is generated from

$$u_{1t} = \rho u_{1t-1} + \varepsilon_{1t} \quad \text{and} \quad u_{2t} = \varepsilon_{2t},$$

$$\text{where} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \sim i.i.d.N(0, \Sigma) \quad \text{with} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}.$$

We set $\rho = 0.1, 0.3, 0.5, 0.7, 0.8, 0.85, 0.9, 0.95, 0.98$ and $T = 100$ and 300, while μ and β are set to be 1 throughout the simulations. The variances σ_{11} and σ_{22} are set to be 1 or 3, and the covariance σ_{21} is chosen such that the correlation between ε_{1t} and ε_{2t} becomes 0.4 or 0.8. The number of replications is 10,000 and all computations are carried out by using the GAUSS matrix language.

We estimate the long-run variances by employing either the kernel method or Andrews and Monahan's (1992) prewhitening method with the specification of a first-order vector autoregression (VAR(1)) for u_t. For the kernel method, we use either the Bartlett or the quadratic spectral (QS) kernel with the bandwidth parameter chosen by either Andrews' (1991) automatic bandwidth selection method or Newey and West's (1994) method. For the prewhitening method, we need to estimate the long-run variances of the regression errors after fitting the VAR(1) model. We employ the same set of the kernel and the bandwidth parameter as the kernel method. Thus, there are eight versions of the long-run variance estimators for the FMR and CCR estimators. Note that since the VAR(1) specification for u_t is correct in our simulations, we can expect the prewhitening method to work better than the kernel method.

With regard to the selection of the lead-lag truncation parameter, we choose K by either the Akaike information criterion (AIC), Bayesian information criterion (BIC), or the general to specific rule as proposed by Ng and Perron (1995) at the 1% or 5% significance level. We set $\bar{K}$, the maximum of K, to be $[12(T/100)^{1/4}]$.

Tables 1 and 2 present the bias and the mean squared error (MSE). The simulation result is summarized as follows:

(i) Both the bias and the MSE become larger for all the estimates as ρ approaches 1.

(ii) All the three efficient methods eliminate the bias of the OLS estimate more or less for all the values of ρ considered in the simulations. However, the effect of the efficient methods when ρ is close to 1 is not pronounced to the extent that it is when ρ is relatively small.

(iii) When ρ is close to 1, the MSE of the efficient estimates is not necessarily smaller than that of the OLS estimate.

(iv) The large variance of ε_{1t} relative to that of ε_{2t} results in the poor performance of all the methods.

(v) The performance of the FMR estimate is similar to that of the CCR estimate, although the bias of the former tends to be slightly smaller than the latter, while the MSE of the former seems to be larger than that of the latter.

(vi) As to the FMR and CCR estimates, the prewhitening method works better than the kernel method unless ρ is very close to 1.

(vii) The performance of the FMR and CCR estimates does not much depend on the selection of the kernel and/or the bandwidth parameter.

(viii) The performance of the DOLS estimates depends considerably on the selection of the lead-lag truncation method.

(i)–(iii) are all related to the case when ρ is close to 1. We observe that all the three efficient methods have some problems when the regression errors are strongly serially correlated.

(iv) implies that the large noise in the regression model affects the estimates in a negative manner. (v) may be expected because both the FMR and CCR methods eliminate the endogeneity bias in the same manner. As we expect, the prewhitening method performs better than the kernel method as stated in (vi). (vii) has occasionally been pointed out in the literature. (viii) is a natural result because the lead-lag truncation parameter K must diverge to infinity as T goes to infinity.

In order to observe the dependence of the DOLS estimate on K in detail, we conduct the simulations for $\rho = 0.5$, 0.8, and 0.9 using a fixed value of K that ranges from 0 to $\bar{K}$. The result is summarized in Figure 1. We observe that while the bias of the estimate is reduced as K grows in the range we considered[4], the MSE tends to be smaller to some extent by the middle of the range but it turns out to be greater for a larger K. From the figure, it appears that there is an optimal K depending on the sample size as well as the strength of a serial correlation. To the best of our knowledge, no optimal method in selecting the lead-lag truncation parameter has been proposed in the literature; this is beyond the scope of this paper.

As we observed, the performance of the estimates depends on ρ and becomes worse when ρ is close to 1. In the next section, we investigate the three efficient estimators as well as the OLS estimator when ρ is close to 1; further, we theoretically explain the finite sample performance observed in this section.

4. Asymptotic Property of the Estimators with Strongly Serially Correlated Errors

This section investigates the asymptotic property of the three efficient estimators as well as the OLS estimator when the cointegrating regression errors are strongly serially correlated. In order to focus on the strength of serial correlation in the regression errors, we consider the model (1) with the following simple structure in the error term:

$$u_{1t} = \rho u_{1t-1} + \varepsilon_{1t}, \quad u_{2t} = \varepsilon_{2t} \tag{11}$$

[4]Of course, the bias will get larger if we consider much larger K.

with $u_{10} = 0$, where $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}']' \sim i.i.d.(0, \Sigma)$ with finite δ-th moment for some $\delta > 4$. This assumption implies that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \varepsilon_t \Rightarrow B(r) = \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix} \tag{12}$$

for $0 \leq r \leq 1$ where $B(\cdot) \sim BM(\Sigma)$.

Before proceeding with the asymptotic analysis, we need to state two assumptions concerned with the kernel method and the lead-lag truncation parameter.

Assumption 1 *(a) The kernel $k(\cdot)$ is continuous at zero, $k(0) = 1$, $\sup_{r \geq 0} |k(r)| \leq \infty$, and $\int_{[0,\infty)} \bar{k}(r)dr < \infty$, where $\bar{k}(r) = \sup_{s \geq r} |k(s)|$. (b) The bandwidth parameter M goes to infinity as $T \to \infty$ and $M = o(T^{1/2})$.*

Assumption 2 *(a) $K = o(T^{1/2})$. (b) $K \sum_{|j|>K} ||\pi_j|| \to 0$.*

Assumption 1 is sufficient to consistently estimate the long-run variances. See Jansson (2002). Assumption 2 is provided by Kejriwal and Perron (2006) and is different from that in Saikkonen (1991) in which it is assumed that (a') $K = o(T^{1/3})$ and (b') $T^{1/2} \sum_{|j|>K} ||\pi_j|| \to 0$. The assumptions (a) and (a') are concerned with the upper bound of K; Kejrwal and Perron (2006) proved that K can increase faster than the rate considered by Saikkonen (1991). The assumptions (b) and (b') provide the lower bound condition on K; the matter of importance is that (b') excludes the case where K is chosen by an information criterion, while (b) is sufficiently general for K to increase at a logarithmic rate and then allow an information criterion for the selection of K.

For the FMR and CCR estimators, we focus on the case where the long-run variances are estimated by the kernel method as follows:

$$\hat{\Omega} = \hat{\Sigma}_u + \hat{\Pi} + \hat{\Pi}' \quad \text{and} \quad \hat{\Lambda} = \hat{\Sigma}_u + \hat{\Pi},$$

$$\text{where} \quad \hat{\Sigma}_u = \hat{\Gamma}(0), \quad \hat{\Pi} = \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right) \hat{\Gamma}(j), \quad \text{and} \quad \hat{\Gamma}(j) = \frac{1}{T} \sum_{t=1}^{T-j} \hat{u}_t \hat{u}_{t+j}'$$

with $\hat{u}_t = [\hat{u}_{1t}, \Delta x_t']'$, where $\hat{u}_{1t}$ is the OLS residual. Note that $\hat{\Gamma}(j)$ is a consistent estimator of $\Gamma(j) = E(u_t u_{t+j}')$.

Using the result in (12) with Assumptions 1 and 2, it is shown that when ρ is fixed, the centered estimators have the same limiting distributions as given by (3) and (10), with $W(\cdot)$ being replaced by $B(\cdot)$. However, when the error term is strongly serially correlated, the assumption of the fixed ρ does not seem appropriate. In this case, we should consider a local-to-unity system such that ρ approaches 1 as T goes to infinity. In the following subsection, we consider two different types of local-to-unity systems.

4.1. Asymptotic property with the T local-to-unity system

First we consider the T local-to-unity system, which is defined as

$$\rho = \rho_T = 1 - \frac{c}{T}, \tag{13}$$

where c is a positive constant. This local-to-unity system has often been assumed in the literature in order to investigate the asymptotic local power of unit root/cointegration tests. See, for example, Phillips (1987), Tanaka (1996), and Saikkonen and Lütkepohl (1999) among others. Note that we usually test for cointegration before estimating cointegrating regression models and that tests for cointegration do not necessarily detect the existence of cointegration with probability 1 even asymptotically if (13) is true. This is because (13) corresponds to the local alternative for tests of cointegration. In other words, we do not always encounter the T local-to-unity system even if (13) is correct. However, it nevertheless appears insightful to investigate the asymptotic property of the estimators with the specification of (13).

Since u_{1t} is a nearly integrated process when ρ is given by (13), it can be shown that

$$\frac{1}{\sqrt{T}} u_{1[Tr]} \Rightarrow J_c(r) \quad \text{for} \quad 0 \le r \le 1, \tag{14}$$

where $J_c(r)$ is the Ornstein-Uhlenbeck process defined as $J_c(r) = \int_0^r e^{-c(r-s)} dB_1(s)$. In this case, the OLS estimator is shown to be inconsistent, which is intuitively explained as follows.

Because of (14) and the continuous mapping theorem (CMT), $\sum_{t=1}^{T} x_t u_{1t}$ is no longer of order T but of order T^2, while $\sum_{t=1}^{T} x_t x_t' = O_p(T^2)$. As a result, it can be seen with no constant term model that $\hat{\beta} = \beta + (\sum_{t=1}^{T} x_t x_t')^{-1} \sum_{t=1}^{T} x_t u_{1t} = \beta + O_p(1)$ and thus the OLS estimator $\hat{\beta}$ is inconsistent. We can further show that the three efficient methods do not work well under the T local-to-unity system. The formal statement is given in the following theorem.

Theorem 1 *Under the T local-to-unity system,*

$$\frac{1}{T} D_T(\hat{\theta}_E - \theta) \Rightarrow H^{-1} h \tag{15}$$

where $\hat{\theta}_E = \hat{\theta}$, $\hat{\beta}_{FMR}$, $\hat{\beta}_{CCR}$, and $\hat{\beta}_{DOLS}$,

$$H = \left[\begin{array}{cc} 1 & \int_0^1 B_2(r)' dr \\ \int_0^1 B_2(r) dr & \int_0^1 B_2(r) B_2(r)' dr \end{array} \right], \quad \textit{and} \quad h = \left[\begin{array}{c} \int_0^1 J_c(r) dr \\ \int_0^1 B_2(r) J_c(r) dr \end{array} \right].$$

Since $T^{-1} D_T = \text{diag}\{T^{-1/2}, I_n\}$, we can see that all the estimators are inconsistent and the estimator of the coefficient associated with a constant term diverges at the rate of $T^{1/2}$.

Theorem 1 shows that the three efficient methods do not work at all and these estimators have the same limiting distribution as the OLS estimator. In other words, the correction terms for the second-order bias are asymptotically negligible for the three efficient estimators. For example, the FMR estimator (4) can be expressed as

$$\begin{aligned} \hat{\theta}_{FMR} &= \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \left\{ \sum_{t=1}^{T} z_t (y_t - \hat{\omega}_{12} \hat{\Omega}_{22}^{-1} u_{2t}) - T \hat{J}^+ \right\} \\ &= \hat{\theta} - \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \left(\sum_{t=1}^{T} z_t u_{2t}' \hat{\Omega}_{22}^{-1} \hat{\omega}_{21} - T \hat{J}^+ \right) \end{aligned}$$

and then, we have

$$\begin{aligned} & \frac{1}{T} D_T(\hat{\theta}_{FMR} - \hat{\theta}) \qquad (16) \\ &= -\left(D_T^{-1} \sum_{t=1}^{T} z_t z_t' D_T^{-1} \right)^{-1} \left(D_T^{-1} \sum_{t=1}^{T} z_t u_{2t}' \hat{\Omega}_{22}^{-1} \frac{1}{T} \hat{\omega}_{21} - \left[\begin{array}{c} 0 \\ \frac{1}{T} (\hat{\lambda}_{21} - \hat{\Lambda}_{22} \hat{\Omega}_{22}^{-1} \hat{\omega}_{21}) \end{array} \right] \right). \end{aligned}$$

Since it is proved in the appendix that $\hat{\omega}_{21}$ and $\hat{\lambda}_{21}$ are $O_p(M)$, we can see that the right-hand side of (16) converges in probability to zero and thus the FMR estimator behaves asymptotically in the same manner as the OLS estimator. The equivalence between the CCR and OLS estimators can be explained in the same way. For the DOLS estimator (7), we can see that

$$\hat{\theta}_{DOLS} = \theta + \left(\sum_{t=K+1}^{T-K} \tilde{z}_t \tilde{z}_t' \right)^{-1} \left(\sum_{t=K+1}^{T-K} \tilde{z}_t \tilde{u}_{1t} \right). \tag{17}$$

Note that $\tilde{\tilde{u}}_{1t}$ does not appear in the above expression, but we have $\tilde{u}_{1t}$ since the relation (8) no longer holds as u_{1t} is a nearly integrated process. Since u_{1t} and x_t dominate w_t, we can show that (17) is asymptotically equivalent to

$$\hat{\theta} = \theta + \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \left(\sum_{t=1}^{T} z_t u_{1t} \right)$$

and the DOLS and OLS estimators thus have the same limiting distribution.

Theorem 1 partially explains why the three methods do not work well and why they cannot effectively eliminate the second-order bias when ρ is very close to 1. However, as the simulation result in the previous section indicates, both the bias and the MSE are still slightly reduced when ρ is moderately close to 1; this result cannot be explained by Theorem 1. Therefore, we need to consider the case where ρ is close to 1 but not as close as under the T local-to-unity system. In the next subsection, we consider the case where ρ approaches 1 at a slower rate than $1/T$.

4.2. Asymptotic property with the N local-to-unity system

This section considers the N local-to-unity system, which is defined as

$$\rho = \rho_N = 1 - \frac{c}{N}, \quad N \to \infty \quad \text{and} \quad \frac{N}{T} \to 0. \tag{18}$$

Under this system, $\rho = \rho_N$ approaches 1 as T goes to infinity; however, the approaching speed is slower than the T local-to-unity system. Note that since ρ_N approaches 1 at a slower rate than ρ_T, tests for cointegration detect the cointegrating relation with an asymptotic probability 1, which implies that we always encounter the N local-to-unity system at least asymptotically if (18) is true. The aim in considering this system is to investigate the behavior of the estimators when the AR coefficient is close to 1 but not too close. As we observed in the previous subsection, the AR coefficient of the T local-to-unity system is very close to 1 and the asymptotic behavior of the estimators is rather different from that when ρ is fixed. The N local-to-unity system may be seen as the bridge between the fixed ρ and the T local-to-unity system. This type of local-to-unity system is also considered by Phillips and Magdalinos (2005, 2007) and Giraitis and Phillips (2006).

The following theorem provides the asymptotic distributions of the estimators.

Theorem 2 *Under the N local-to-unity system,*

$$\begin{aligned}
\frac{1}{N}D_T(\hat{\theta}-\theta) &\Rightarrow H^{-1}h_{OLS}, && (19)\\
\frac{1}{N}D_T(\hat{\theta}_{FM}-\theta) &\Rightarrow H^{-1}h_{FMR}, && (20)\\
\frac{1}{N}D_T(\hat{\theta}_{CCR}-\theta) &\Rightarrow H^{-1}h_{FMR}, && (21)\\
\frac{1}{N}D_T(\hat{\theta}_{DOLS}-\theta) &\Rightarrow H^{-1}h_{DOLS}, && (22)
\end{aligned}$$

where H is defined in Theorem 1 and

$$h_{OLS} = \begin{bmatrix} \frac{1}{c}B_1(1) \\ \frac{1}{c}\left(\int_0^1 B_2(r)dB_1(r)+\sigma_{21}\right) \end{bmatrix},$$

$$h_{FMR} = \begin{cases}
\begin{bmatrix} \frac{1}{c}B_{1\cdot 2}(1) \\ \frac{1}{c}\int_0^1 B_2(r)dB_{1\cdot 2}(r) \end{bmatrix}, & \frac{M}{N}\to\infty,\\
\begin{bmatrix} \frac{1}{c}B_1(1) - B_2'(1)\Sigma_{22}^{-1}\sigma_{21}d_f\int_0^\infty k(r)e^{-cd_f r}dr \\ \Big[\frac{1}{c}\int_0^1 B_2(r)dB_1(r) \\ -\int_0^1 B_2(r)dB_2'(r)\Sigma_{22}^{-1}\sigma_{21}d_f\int_0^\infty k(r)e^{-cd_f r}dr \\ +\left(\frac{1}{c}-d_f\int_0^\infty k(r)e^{-cd_f r}dr\right)\sigma_{21}\Big] \end{bmatrix}, & \frac{M}{N}\to d_f,\\
\begin{bmatrix} \frac{1}{c}B_1(1) \\ \frac{1}{c}\left(\int_0^1 B_2(r)dB_1(r)+\sigma_{21}\right) \end{bmatrix}, & \frac{M}{N}\to 0,
\end{cases}$$

$$h_{DOLS} \;=\; \begin{cases} \begin{bmatrix} \frac{1}{c}B_{1\cdot 2}(1) \\ \frac{1}{c}\int_0^1 B_2(r)dB_{1\cdot 2}(r) \end{bmatrix}, & \frac{K}{N} \to \infty, \\ \begin{bmatrix} \frac{1}{c}B_1(1) - (1-e^{-cd_d})\sigma_{12}\Sigma_{22}^{-1}B_2(1) \\ \left[\frac{1}{c}\left(\int_0^1 B_2(r)dB_1(r) \right.\right. \\ \left.\left. -(1-e^{-cd_d})\int_0^1 B_2(r)dB_2'(r)\Sigma_{22}^{-1}\sigma_{21} + e^{-cd_d}\sigma_{21}\right)\right] \end{bmatrix}, & \frac{K}{N} \to d_d, \\ \begin{bmatrix} \frac{1}{c}B_1(1) \\ \frac{1}{c}\left(\int_0^1 B_2(r)dB_1(r) + \sigma_{21}\right) \end{bmatrix}, & \frac{K}{N} \to 0, \end{cases}$$

with d_f and d_d being fixed positive values and $B_{1\cdot 2}(r) = B_1(r) - \sigma_{12}\Sigma_{22}^{-1}B_2(r)$.

Note that all the estimators are consistent under the N local-to-unity system; however, they are not T-consistent as in the case where ρ is fixed but the convergence rate is slower than T.

From Theorem 2, we observe that the three efficient estimators do not have the second-order bias when N is sufficiently slow as compared with the bandwidth parameter M or the lead-lag truncation parameter K. This implies that when ρ approaches 1 slowly or ρ is sufficiently away from 1, the FMR, CCR, and DOLS estimators are efficient as compared with the OLS estimator. On the other hand, when ρ approaches 1 rapidly or when ρ is very close to 1, they have the same asymptotic distributions as the OLS estimator and hence suffer from the second-order bias. For the intermediate case where N is of the same order as M or K, the limiting distributions of the three efficient estimators become complicated. In this case, the second-order bias is partially eliminated by the efficient methods. For example, the endogeneity bias is partially eliminated from the FMR and CCR estimators by observing the term

$$\frac{1}{c}dB_1(r) - dB_2'(r)\Sigma_{22}^{-1}\sigma_{21}d_f\int_0^\infty k(r)e^{-cd_f r}dr,$$

while the noncentrality is adjusted by the term

$$\left(\frac{1}{c} - d_f\int_0^\infty k(r)e^{-cd_f r}dr\right)\sigma_{21}.$$

Similar effects can be observed for the limiting distribution of the DOLS estimator when $K/N \to d_d$. Thus, Theorem 2 implies that when ρ is relatively away from 1, the three efficient estimators work well as compared with the OLS estimator; however, as ρ approaches 1, the difference between the efficient estimators and the OLS estimator reduces, and eventually, when ρ is sufficiently close to 1, the difference becomes negligible. This is consistent with the finite sample behavior of the estimators observed in Section 3. In other words, the N local-to-unity system can adequately explain the finite sample behavior of the estimators when ρ is moderately close to 1.

Finally, we demonstrate the probability density functions (pdf) of the distributions provided in Theorem 2. Figure 2 illustrates the pdfs[5] for $\sigma_{21} = 0.4$ and 0.8, $c = 1/2$, 1, 2, and 3, and $d_d = d_f = 1$. They are obtained from 100,000 replications from the distribution of the discrete approximation based on 2,000 steps to the limiting distribution provided in Theorem 2. We can observe that the limiting distribution for a slow N is centered at and symmetric around the origin, while the limiting distribution of the OLS estimator is shifted and skewed toward the right-hand side. In addition, we observe that the second-order bias is partially eliminated by the efficient methods by observing the limiting distributions corresponding to the cases where $M/N = 1$ and $K/N = 1$. Overall, the second-order bias of the OLS estimator increases for a larger σ_{21} and a smaller c.

5. Conclusion

In this paper, we theoretically investigate the three efficient estimators for cointegrating regression models: the FMR, CCR, and DOLS estimators. We first demonstrate that these efficient methods do not necessarily work well in finite samples when ρ is close to 1. Next, we show that the three efficient methods work as poorly as the OLS method when the regression errors are generated from the T local-to-unity system. Furthermore, by using the N local-to-unity system, the three efficient methods are shown to estimate the regression coefficient more efficiently than the OLS method when the regression errors are moderately serially correlated; however, as the correlation becomes stronger, the effect of the efficient methods is

[5] These densities are drawn for the range of 1% to 99% points by the kernel method with a Gaussian kernel. The smoothing parameter, h, is decided by equation (3.31) in Silverman (1986): $h = 0.9AT^{-1/5}$ where $A = \min(\text{standard deviation, interquartile range}/1.34)$.

weakened, and eventually, all the estimators considered in this paper become asymptotically equivalent when the correlation is sufficiently strong. This analytical investigation can adequately explain the poor finite sample behavior of the three efficient estimators when the regression errors are serially correlated.

Appendix

In the appendix, we denote some constants that are independent of T and the subscript j as C_1, C_2, $\cdots$.

Proof of Theorem 1: Noting that

$$\frac{1}{T}D_T(\hat{\theta}-\theta) = \left[\begin{array}{cc} 1 & \frac{1}{T\sqrt{T}}\sum_{t=1}^T x_t' \\ \frac{1}{T\sqrt{T}}\sum_{t=1}^T x_t & \frac{1}{T^2}\sum_{t=1}^T x_t x_t' \end{array}\right]\left[\begin{array}{c} \frac{1}{T\sqrt{T}}\sum_{t=1}^T u_{1t} \\ \frac{1}{T^2}\sum_{t=1}^T x_t u_{1t} \end{array}\right],$$

we obtain (15) for the OLS estimator $\hat{\theta}$ using (12), (14), and the CMT.

For the FMR estimator, it is sufficient to prove that $\hat{\omega}_{21}$ and $\hat{\lambda}_{21}$ are $O_p(M)$ as explained after Theorem 1, because the centered FMR estimator is expressed as (16) while $D_T^{-1}\sum_{t=1}^T z_t z_t' D_T^{-1} \Rightarrow H$ and $D_T^{-1}\sum_{t=1}^T z_t u_{2t}' \Rightarrow [B_2'(1), (\int_0^1 B_2(r)dB_2'(r) + \Lambda_{22})']'$. Note that the one sided version of the long-run variance estimator is defined as

$$\hat{\lambda}_{21} = \sum_{j=0}^{T-1} k\left(\frac{j}{M}\right)\frac{1}{T}\sum_{t=1}^{T-j} u_{2t}\hat{u}_{1t+j}.$$

Since $T^{-1}\sum_{t=1}^{T-j} u_{2t}\hat{u}_{1t+j}$ is shown to be $O_p(1)$ uniformly over j and $M^{-1}\sum_{j=0}^{T-1} k\left(\frac{j}{M}\right) \to \int_0^\infty k(r)dr$, we can see that $\hat{\lambda}_{21}$ is $O_p(M)$. Since $\hat{\omega}_{21}$ is similarly shown to be $O_p(M)$ and $M/T \to 0$ by Assumption 1(b), we obtain the theorem for the FMR estimator.

For the CCR estimator, we have

$$\begin{aligned} \frac{1}{T}D_T(\hat{\theta}_{CCR}-\theta) &= \left(D_T^{-1}\sum_{t=1}^T z_t^* z_t^{*\prime} D_T^{-1}\right)^{-1}\frac{1}{T}D_T^{-1}\sum_{t=1}^T z_t^* u_{1t}^* \\ &= \left[\begin{array}{cc} 1 & \frac{1}{T\sqrt{T}}\sum_{t=1}^T x_t^{*\prime} \\ \frac{1}{T\sqrt{T}}\sum_{t=1}^T x_t^* & \frac{1}{T^2}\sum_{t=1}^T x_t^* x_t^{*\prime} \end{array}\right]^{-1}\left[\begin{array}{c} \frac{1}{T\sqrt{T}}\sum_{t=1}^T u_{1t}^* \\ \frac{1}{T^2}\sum_{t=1}^T x_t^* u_{1t}^* \end{array}\right] \end{aligned}$$

where $u_{1t}^* = u_{1t} - (\hat{\beta}-\beta)'\hat{\Lambda}_2\hat{\Sigma}_u^{-1}\hat{u}_t - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}u_{2t}$, $\hat{\beta}$ being the OLS estimator of β.

In order to derive the limiting distribution, we need to obtain the order of $\hat{\Lambda}_2\hat{\Sigma}_u^{-1}$. Since

$$\begin{aligned} \frac{1}{\sqrt{T}}\hat{u}_{1[Tr]} &= \frac{1}{\sqrt{T}}u_{1[Tr]} - \sqrt{T}z_{[Tr]}'D_T^{-1}\cdot\frac{1}{T}D_T(\hat{\theta}-\theta) \\ &\Rightarrow J_c(r) - [1, B_2'(r)]H^{-1}h \equiv \bar{J}_c(r), \quad \text{say,} \end{aligned} \qquad (23)$$

we have $T^{-1}\hat{\sigma}_{u11} \Rightarrow \int_0^1 \bar{J}_c^2(r)dr$ and $\hat{\sigma}_{u21} = O_p(1)$, the latter of which can be shown as Theorem 4.1 of Hansen (1992). Thus, since $\hat{\lambda}_{21} = O_p(M)$, we obtain

$$\begin{aligned}
\hat{\Lambda}_2\hat{\Sigma}_u^{-1}D_T^{-1}T &= \hat{\Lambda}_2 D_T \frac{1}{T}\left(\frac{1}{T^2}D_T\hat{\Sigma}_u D_T\right)^{-1} \\
&= \left[\frac{1}{\sqrt{T}}\hat{\lambda}_{21}, \hat{\Lambda}_{22}\right]\begin{bmatrix} \frac{1}{T}\hat{\sigma}_{u11} & \frac{1}{\sqrt{T}}\hat{\sigma}_{u12} \\ \frac{1}{\sqrt{T}}\hat{\sigma}_{u21} & \hat{\Sigma}_{u22}\end{bmatrix}^{-1} \\
&\Rightarrow [0, \Lambda_{22}]\begin{bmatrix} (\int_0^1 \bar{J}_c^2(r)dr)^{-1} & 0 \\ 0 & \Sigma_{u22}^{-1}\end{bmatrix} = [0, I_n] \qquad (24)
\end{aligned}$$

where the last equality holds because $\Lambda_{22} = \Omega_{22} = \Sigma_{u22}$ in our model. Then, noting that $\sum_{t=1}^T \hat{u}_{1t} = 0$ because $\hat{u}_{1t}$ is the regression residual and that $T^{-1/2}\sum_{t=1}^T u_{2t} \Rightarrow B_2(1)$, we can see that

$$\begin{aligned}
\frac{1}{T\sqrt{T}}\sum_{t=1}^T x_t^* &= \frac{1}{T\sqrt{T}}\sum_{t=1}^T x_t - \hat{\Lambda}_2\hat{\Sigma}_u^{-1}D_T^{-1}T \cdot \frac{1}{T^2\sqrt{T}}D_T\sum_{t=1}^T \hat{u}_t \\
&= \frac{1}{T\sqrt{T}}\sum_{t=1}^T x_t + O_p\left(\frac{1}{T}\right) \\
&\Rightarrow \int_0^1 B_2(r)dr. \qquad (25)
\end{aligned}$$

Similarly, we have

$$\frac{1}{T^2}\sum_{t=1}^T x_t^* x_t^{*\prime} = \frac{1}{T^2}\sum_{t=1}^T x_t x_t' + o_p(1) \Rightarrow \int_0^1 B_2(r)B_2'(r)dr. \qquad (26)$$

On the other hand, since $\hat{\beta} = O_p(1)$ and $\hat{\omega}_{12} = O_p(M)$, using (24), we can see that

$$\begin{aligned}
\frac{1}{T\sqrt{T}}\sum_{t=1}^T u_{1t}^* &= \frac{1}{T\sqrt{T}}\sum_{t=1}^T u_{1t} - (\hat{\beta}-\beta)'\hat{\Lambda}_2\hat{\Sigma}_u^{-1}D_T^{-1}T \cdot \frac{1}{T^2\sqrt{T}}D_T\sum_{t=1}^T \hat{u}_t \\
&\quad -\frac{1}{T}\hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T u_{2t} \\
&= \frac{1}{T\sqrt{T}}\sum_{t=1}^T u_{1t} + O_p\left(\frac{1}{T}\right) + O_p\left(\frac{M}{T}\right) \\
&\Rightarrow \int_0^1 J_c(r)dr \qquad (27)
\end{aligned}$$

and

$$
\begin{aligned}
\frac{1}{T^2}\sum_{t=1}^{T} x_t^* u_{1t}^* &= \frac{1}{T^2}\sum_{t=1}^{T} x_t u_{1t} - \frac{1}{T^3}\sum_{t=1}^{T} x_t \hat{u}_t' D_T \cdot T D_T^{-1}\hat{\Sigma}_u^{-1}\hat{\Lambda}_2'(\hat{\beta}-\beta) \\
&\quad -\frac{1}{T}\sum_{t=1}^{T} x_t u_{2t}'\hat{\Omega}_{22}^{-1}\frac{1}{T}\hat{\omega}_{21} - \hat{\Lambda}_2\hat{\Sigma}_u^{-1}D_T^{-1}T \cdot D_T\frac{1}{T^3}\sum_{t=1}^{T}\hat{u}_t u_{1t} \\
&\quad +\hat{\Lambda}_2\hat{\Sigma}_u^{-1}D_T^{-1}T \cdot \frac{1}{T^4}D_T\sum_{t=1}^{T}\hat{u}_t\hat{u}_t' D_T \cdot T D_T^{-1}\hat{\Sigma}_u^{-1}\hat{\Lambda}_2'(\hat{\beta}-\beta) \\
&\quad +\hat{\Lambda}_2\hat{\Sigma}_u^{-1}D_T^{-1}T \cdot \frac{1}{T^2}D_T\sum_{t=1}^{T}\hat{u}_t u_{2t}'\hat{\Omega}_{22}^{-1}\frac{1}{T}\hat{\omega}_{21} \\
&= \frac{1}{T^2}\sum_{t=1}^{T} x_t u_{1t} + O_p\left(\frac{1}{T}\right) + O_p\left(\frac{M}{T}\right) + o_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{T}\right) + O_p\left(\frac{M}{T}\right) \\
&\Rightarrow \int_0^1 B_2(r)J_c(r)dr. \qquad (28)
\end{aligned}
$$

From (25) – (28), we obtain the result for the CCR estimator.

For the DOLS estimator, we note that the relation in (8) no longer holds. From (17) and the inverse formula of the partitioned matrix, the DOLS estimator can be expressed as

$$
\frac{1}{T}D_T(\hat{\theta}_{DOLS}-\theta) = \left(D_T^{-1}\sum_{t=K+1}^{T-K} z_t z_t' D_T^{-1} - G_1 G_2^{-1} G_1'\right)^{-1}\left(\frac{1}{T}D_T^{-1}\sum_{t=K+1}^{T-K} z_t u_{1t} - G_1 G_2^{-1} G_3\right)
$$

$$
\text{where}\quad G_1 = \frac{1}{\sqrt{T}}D_T^{-1}\sum_{t=K+1}^{T-K} z_t w_t', \quad G_2 = \frac{1}{T}\sum_{t=K+1}^{T-K} w_t w_t', \quad \text{and} \quad G_3 = \frac{1}{T\sqrt{T}}\sum_{t=K+1}^{T-K} w_t u_{1t}.
$$

From this expression, we can see that the result for the DOLS estimator is obtained if we show that

$$
||G_1 G_2^{-1} G_1'|| = o_p(1) \quad \text{and} \quad ||G_1 G_2^{-1} G_3|| = o_p(1)
$$

where $\|\cdot\|$ is a matrix norm defined by $\|A\| = (tr(A'A))^{1/2}$ for a given matrix A. Note that $||A_1 A_2|| \leq ||A_1|| ||A_2||_1 \leq ||A_1|| ||A_2||$ because of Lemma A1 in Saikkonen (1991) and $||A_1||_1^2 \leq ||A^2||$ for given matrices A_1 and A_2, where $||A||_1 = \sup\{||Ax|| : ||x|| \leq 1\}$. Since $||G_2^{-1}||_1$ is shown to be $O_p(1)$ by Saikkonen (1991) and Kejriwal and Perron (2006), it is sufficient to show that $||G_1|| = o_p(1)$ and $||G_3|| = o_p(1)$. Here, note that

$$\begin{aligned}
\|G_1\|^2 &= \frac{1}{T}\left[\left\|\frac{1}{\sqrt{T}}\sum_{t=K+1}^{T-K} w_t'\right\|^2 + \left\|\frac{1}{T}\sum_{t=K+1}^{T-K} x_t w_t'\right\|^2\right] \\
&= \frac{1}{T}\left[\sum_{j=-K}^{K}\left\|\frac{1}{\sqrt{T}}\sum_{t=K+1}^{T-K} u_{2,t-j}'\right\|^2 + \sum_{j=-K}^{K}\left\|\frac{1}{T}\sum_{t=K+1}^{T-K} x_t u_{2,t-j}'\right\|^2\right] \\
&= O_p\left(\frac{K}{T}\right), \qquad (29) \\
||G_3||^2 &= \frac{1}{T^3}\sum_{j=-K}^{K}\left\|\sum_{t=K+1}^{T-K} u_{2t-j}u_{1t}\right\|^2 \\
&\leq \frac{1}{T^3}\sup_t ||u_{1t}||^2 \sum_{j=-K}^{K}\left\|\sum_{t=K+1}^{T-K} u_{2t-j}\right\|^2 = O_p\left(\frac{K}{T}\right).
\end{aligned}$$

Hence, $\|G_1\| = o_p(1)$ and $||G_3|| = o_p(1)$.■

Proof of Theorem 2: First, we prove the following lemma.

Lemma A.1 *For* $\rho = \rho_N = 1 - c/N$,

$$(a) \quad \frac{1}{NT}\sum_{t=1}^{T} u_{1t}^2 \xrightarrow{p} \frac{1}{2c}\sigma_{11},$$

$$(b) \quad \frac{1}{N\sqrt{T}}\sum_{t=1}^{T} u_{1t} \Rightarrow \frac{1}{c}B_1(1),$$

$$(c) \quad \frac{1}{NT}\sum_{t=1}^{T} x_t u_{1t} \Rightarrow \frac{1}{c}\left(\int_0^1 B_2(r)dB_1(r) + \sigma_{21}\right).$$

Proof of Lemma A.1: (a) Since $u_{1t} = \sum_{l=1}^{t}\rho^{t-l}\varepsilon_{1l}$, we have

$$\sum_{t=1}^{T} u_{1t}^2 = \sum_{t=1}^{T}\sum_{l=1}^{t}\rho^{2(t-l)}\varepsilon_{1l}^2 + 2\sum_{t=2}^{T}\sum_{l=1}^{t-1}\sum_{k=l+1}^{t}\rho^{2t-k-l}\varepsilon_{1k}\varepsilon_{1l}. \qquad (30)$$

The first term on the right-hand side of (30) can be expressed as

$$\sum_{t=1}^{T}\sum_{l=1}^{t}\rho^{2(t-l)}\varepsilon_{1l}^2$$

$$
\begin{array}{ccccccccc}
= & \varepsilon_{11}^2 & & & & & & & \\
+ & \rho^2\varepsilon_{11}^2 & + & \varepsilon_{12}^2 & & & & & \\
+ & \rho^4\varepsilon_{11}^2 & + & \rho^2\varepsilon_{12}^2 & + & \varepsilon_{13}^2 & & & \\
 & \vdots & & \vdots & & \vdots & \ddots & & \\
+ & \rho^{2(T-1)}\varepsilon_{11}^2 & + & \rho^{2(T-2)}\varepsilon_{12}^2 & + & \rho^{2(T-3)}\varepsilon_{13}^2 & +\cdots+ & \varepsilon_{1T}^2 &
\end{array} \tag{31}
$$

$$
\begin{aligned}
&= \sum_{t=1}^{T}\left(1+\rho^2+\cdots+\rho^{2(T-t)}\right)\varepsilon_{1t}^2 \\
&= \frac{1}{1-\rho^2}\sum_{t=1}^{T}\varepsilon_{1t}^2 - \frac{1}{1-\rho^2}\sum_{t=1}^{T}\rho^{2(T-t+1)}\varepsilon_{1t}^2 \\
&= \frac{1}{1-\rho^2}\sum_{t=1}^{T}\varepsilon_{1t}^2 - \frac{1}{1-\rho^2}\sum_{t=1}^{T}\rho^{2(T-t+1)}(\varepsilon_{1t}^2-\sigma_{11}) - \frac{\sigma_{11}}{1-\rho^2}\sum_{t=1}^{T}\rho^{2(T-t+1)}, \qquad (32)
\end{aligned}
$$

where the second equality is obtained by collecting the coefficients associated with each t. Since $1-\rho^2 = 2c/N - c^2/N^2$, the first term of (32) divided by NT converges in probability to $\sigma_{11}/(2c)$ by the weak law of large numbers (WLLN). The second term of (32) has mean zero and its variance is given by

$$
\begin{aligned}
V\left(\frac{1}{1-\rho^2}\sum_{t=1}^{T}\rho^{2(T-t+1)}(\varepsilon_{1t}^2-\sigma_{11})\right) &= \frac{1}{(1-\rho^2)^2}\sum_{t=1}^{T}\rho^{4(T-t+1)}E\left((\varepsilon_{1t}^2-\sigma_{11})^2\right) \\
&= \frac{\rho^4-\rho^{4(T+1)}}{(1-\rho^2)^2(1-\rho^4)}E\left((\varepsilon_{1t}^2-\sigma_{11})^2\right) = O(N^3).
\end{aligned}
$$

Hence, the second term of (32) is $O_p(N^{3/2})$. Since the third term of (32) is easily shown to be $O(N^2)$, we can see that the first term on the right-hand side of (32) dominates the other terms and thus converges in probability as follows:

$$
\frac{1}{NT}\sum_{t=1}^{T}\sum_{l=1}^{t}\rho^{2(t-l)}\varepsilon_{1l}^2 \xrightarrow{p} \frac{\sigma_{11}}{2c}.
$$

Thus, (a) is established if we prove that the second term on the right-hand side of (30) divided by NT converges in probability to zero. After some algebra, it is shown that

$$
\begin{aligned}
\sum_{t=2}^{T}\sum_{l=1}^{t-1}\sum_{k=l+1}^{t}\rho^{2t-k-l}\varepsilon_{1k}\varepsilon_{1l} &= \sum_{l=1}^{T-1}\sum_{k=l+1}^{T}(\rho^{k-l}+\rho^{k-l+2}+\cdots+\rho^{2T-k-l})\varepsilon_{1k}\varepsilon_{1l} \\
&= \frac{1}{1-\rho^2}\sum_{l=1}^{T-1}\sum_{k=l+1}^{T}(\rho^{k-l}-\rho^{2T-k-l+2})\varepsilon_{1k}\varepsilon_{1l}. \qquad (33)
\end{aligned}
$$

Since (33) has mean zero and its variance is given by

$$V\left(\frac{1}{1-\rho^2}\sum_{l=1}^{T-1}\sum_{k=l+1}^{T}(\rho^{k-l}-\rho^{2T-k-l+2})\varepsilon_{1k}\varepsilon_{1l}\right)=\frac{\sigma_{11}^4}{(1-\rho^2)^2}\sum_{l=1}^{T-1}\sum_{k=l+1}^{T}(\rho^{k-l}-\rho^{2T-k-l+2})^2,$$

which is shown to be $O(N^3T)$. Thus, (33) is $O_p(N^{3/2}T^{1/2})$ and hence the second term on the right-hand side of (30) divided by NT is $O_p(N^{1/2}/T^{1/2})=o_p(1)$.

(b) Using $u_{1t}=\sum_{l=1}^{t}\rho^{t-l}\varepsilon_{1l}$ and $1-\rho=c/N$, we obtain

$$\begin{aligned}\frac{1}{N\sqrt{T}}\sum_{t=1}^{T}u_{1t} &= \frac{1}{N\sqrt{T}}\sum_{t=1}^{T}\sum_{l=1}^{t}\rho^{t-l}\varepsilon_{1l}\\ &= \frac{1}{N\sqrt{T}}\sum_{t=1}^{T}\left(1+\rho+\cdots+\rho^{T-t}\right)\varepsilon_{1t}\\ &= \frac{1}{(1-\rho)N\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{1t}-\frac{1}{(1-\rho)N\sqrt{T}}\sum_{t=1}^{T}\rho^{T-t+1}\varepsilon_{1t} \qquad (34)\\ &= \frac{1}{c\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{1t}+O_p\left(\sqrt{\frac{N}{T}}\right) \qquad (35)\\ &\Rightarrow \frac{1}{c}B_1(1),\end{aligned}$$

where (35) holds because the second term of (34) has mean zero and its variance is shown to be $O(N/T)$.

(c) Using the identity $(1/N)u_{1t-1}=(1/c)\varepsilon_{1t}+(1/c)(u_{1t-1}-u_{1t})$, we have

$$\frac{1}{NT}\sum_{t=1}^{T}x_tu_{1t}=\frac{1}{cT}\sum_{t=1}^{T}x_t\varepsilon_{1t+1}+\frac{1}{cT}\sum_{t=1}^{T}x_t(u_{1t}-u_{1t+1}).$$

The first term on the right-hand side weakly converges to $c^{-1}\int_0^1 B_2(r)dB_1(r)$, while the second term is expressed as

$$\begin{aligned}\frac{1}{cT}\sum_{t=1}^{T}x_t(u_{1t}-u_{1t+1}) &= \frac{1}{cT}\sum_{t=1}^{T}\{(x_t-x_{t-1})u_{1t}+(x_{t-1}u_{1t}-x_tu_{1t+1})\}\\ &= \frac{1}{cT}\sum_{t=1}^{T}\varepsilon_{2t}u_{1t}+\frac{1}{cT}(x_0u_{11}-x_Tu_{1T+1}). \qquad (36)\end{aligned}$$

The first term of (36) converges in probability to σ_{21}/c by the WLLN because it has mean σ_{21}/c and its variance is shown to converge to zero. On the other hand, the second term of (36) converges in probability to zero because $T^{-1/2}\sup_t |x_t| = O_p(1)$ while $T^{-1/2}u_{1t}$ can be shown to converge in probability to zero uniformly over t. We then obtain (c).□

Using Lemma A.1, the limiting distribution of the OLS estimator is obtained as

$$
\begin{aligned}
\frac{1}{N}D_T(\hat{\theta}-\theta) &= \begin{bmatrix} 1 & \frac{1}{T\sqrt{T}}\sum_{t=1}^{T} x_t' \\ \frac{1}{T\sqrt{T}}\sum_{t=1}^{T} x_t & \frac{1}{T^2}\sum_{t=1}^{T} x_t x_t' \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{N\sqrt{T}}\sum_{t=1}^{T} u_{1t} \\ \frac{1}{NT}\sum_{t=1}^{T} x_t u_{1t} \end{bmatrix} \\
&\Rightarrow \begin{bmatrix} 1 & \int_0^1 B_2'(r)dr \\ \int_0^1 B_2(r)dr & \int_0^1 B_2(r)B_2'(r)dr \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{c}B_1(1) \\ \frac{1}{c}\left(\int_0^1 B_2(r)dB_1(r)+\sigma_{21}\right) \end{bmatrix}.
\end{aligned}
$$

Next, we present the following lemma, which is used to derive the limiting distributions of the FMR and CCR estimators.

Lemma A.2 *For $\rho = \rho_N = 1 - c/N$,*

$$(a) \quad \frac{1}{N}\hat{\sigma}_{11} = \frac{1}{NT}\sum_{t=1}^{T}\hat{u}_{1t}^2 \xrightarrow{p} \frac{1}{2c}\sigma_{11},$$

$$(b) \quad \hat{\sigma}_{21} = \frac{1}{T}\sum_{t=1}^{T} u_{2t}\hat{u}_{1t} \xrightarrow{p} \sigma_{21},$$

$$(c) \quad \frac{1}{N}\hat{\lambda}_{21} \xrightarrow{p} \bar{\lambda}_{21} \equiv \begin{cases} \frac{1}{c}\sigma_{21}, & \frac{M}{N} \to \infty, \\ \sigma_{21}d_f\int_0^\infty k(r)e^{-cd_f r}dr, & \frac{M}{N} \to d_f, \\ 0, & \frac{M}{N} \to 0, \end{cases}$$

$$(d) \quad \frac{1}{N}\hat{\omega}_{21} \xrightarrow{p} \bar{\lambda}_{21}.$$

Proof of Lemma A.2: (a) Since $\hat{u}_{1t} = u_{1t} - (\hat{\theta}-\theta)'z_t$ and $(1/N)D_T(\hat{\theta}-\theta) = O_p(1)$, we can see that

$$
\begin{aligned}
\frac{1}{NT}\sum_{t=1}^{T}\hat{u}_{1t}^2 &= \frac{1}{NT}\sum_{t=1}^{T}u_{1t}^2 - 2(\hat{\theta}-\theta)'D_T\frac{1}{N}\cdot\frac{1}{T}D_T^{-1}\sum_{t=1}^{T}z_t u_{1t} \\
&\quad +(\hat{\theta}-\theta)'D_T\frac{1}{N}\cdot\frac{N}{T}D_T^{-1}\sum_{t=1}^{T}z_t z_t' D_T^{-1}\frac{1}{N}D_T(\hat{\theta}-\theta) \\
&= \frac{1}{NT}\sum_{t=1}^{T}u_{1t}^2 + O_p\left(\frac{N}{T}\right) \xrightarrow{p} \frac{1}{2c}\sigma_{11},
\end{aligned}
$$

where the last convergence is obtained from Lemma A.1(a).

(b), (c) First, we note that

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T-j} u_{2t}\hat{u}_{1t+j} &= \frac{1}{T}\sum_{t=1}^{T-j} u_{2t}u_{1t+j} - \frac{N}{T}\sum_{t=1}^{T-j} u_{2t}z_{t+j}' D_T^{-1} \cdot \frac{1}{N} D_T(\hat{\theta}-\theta) \\
&= \frac{1}{T}\sum_{t=1}^{T-j} u_{2t}u_{1t+j} + O_p\left(\frac{N}{T}\right) \qquad (37)
\end{aligned}$$

uniformly over j. Since $u_{1t+j} = \sum_{l=1}^{t+j} \rho^{t+j-l}\varepsilon_{1l}$, the first term of (37) can be expressed as

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T-j} u_{2t}u_{1t+j} &= \frac{\rho^j}{T}\sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t} + \frac{1}{T}\sum_{t=1}^{T-j} u_{2t}\sum_{l=1}^{t-1}\rho^{t+j-l}\varepsilon_{1l} + \frac{1}{T}\sum_{t=1}^{T-j} u_{2t}\sum_{l=t+1}^{t+j}\rho^{t+j-l}\varepsilon_{1l} \\
&= \frac{\rho^j}{T}\sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t} + L_1 + L_2, \quad \text{say.} \qquad (38)
\end{aligned}$$

After some algebra, it can be shown that $E(L_1) = 0$ and $E(L_2) = 0$ while

$$V(L_1) \le C_1\frac{N}{T} \quad \text{and} \quad V(L_2) \le C_2\frac{N}{T}.$$

This implies that both L_1 and L_2 are $O_p(N^{1/2}/T^{1/2})$ uniformly over $j \ge 0$. Then, (b) is obtained from (38) for $j = 0$ by the WLLN.

For (c), we obtain, from (37) and (38),

$$\begin{aligned}
\frac{1}{N}\hat{\lambda}_{21} &= \frac{1}{N}\sum_{j=0}^{T-1} k\left(\frac{j}{M}\right)\frac{1}{T}\sum_{t=1}^{T-j} u_{2t}\hat{u}_{1t+j} \\
&= \frac{1}{N}\sum_{j=0}^{T-1} k\left(\frac{j}{M}\right)\frac{1}{T}\sum_{t=1}^{T-j} u_{2t}u_{1t+j} + O_p\left(\frac{M}{T}\right) \\
&= \frac{1}{N}\left(\sum_{j=0}^{N^*} + \sum_{j=N^*+1}^{T-1}\right) k\left(\frac{j}{M}\right)\frac{\rho^j}{T}\sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t} + O_p\left(\frac{M}{\sqrt{NT}}\right) + O_p\left(\frac{M}{T}\right) \qquad (39)
\end{aligned}$$

where N^* satisfies $N^*/N \to \infty$ and $N^*/T \to 0$. Since $T^{-1}\sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t}$ is shown to be $O_p(1)$ uniformly over $j > N^*$ and $k(\cdot) < C_3$ for some constant C_3, the second summation in the first term of (39) becomes

$$\left|\frac{1}{N}\sum_{j=N^*+1}^{T-1} k\left(\frac{j}{M}\right)\frac{\rho^j}{T}\sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t}\right| \le \frac{C_3}{N}\sum_{j=N^*+1}^{T-1}\rho^j \times O_p(1)$$

$$= \frac{\rho^{N^*+1} - \rho^T}{N(1-\rho)} \times O_p(1),$$

which converges in probability to zero because $N(1-\rho) = c$. Further,

$$\rho^{N^*+1} = \left(1 - \frac{c}{N}\right)^{N^*+1} \to 0 \quad \text{and} \quad \rho^T = \left(1 - \frac{c}{N}\right)^T \to 0$$

since $c > 0$, $N^*/N \to \infty$, and $T/N \to \infty$. In the following, we derive the limit of the first summation in the first term of (39) depending on the rate of N.

When $M/N \to \infty$, we can choose N^* such that $N^*/M \to 0$ while $N^*/N \to \infty$. Then, since $k(0) = 1$, $k(\cdot)$ is continuous at zero, $j/M \to 0$ over $0 \le j \le N^*$, and $T^{-1} \sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t}$ converges in probability to σ_{21} uniformly over $0 \le j \le N^*$, we observe that

$$\begin{aligned}
\frac{1}{N}\sum_{j=0}^{N^*} k\left(\frac{j}{M}\right) \frac{\rho^j}{T} \sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t} &= (\sigma_{21} + o_p(1)) \frac{1}{N} \sum_{j=0}^{N^*} (1 + o(1))\rho^j \\
&\xrightarrow{p} \sigma_{21} \lim_{N\to\infty} \frac{1 - \rho^{N^*+1}}{N(1-\rho)} = \frac{1}{c}\sigma_{21}.
\end{aligned}$$

Thus, we obtain (c) when $M/N \to \infty$.

When $M/N \to d_f$, N^*/M must go to infinity. Then, we have

$$\begin{aligned}
\frac{1}{N}\sum_{j=0}^{N^*} k\left(\frac{j}{M}\right) \frac{\rho^j}{T} \sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t} &= (\sigma_{21} + o_p(1)) \frac{1}{N} \sum_{j=0}^{N^*} k\left(\frac{j}{M}\right) \rho^j \\
&= (\sigma_{21} + o_p(1)) \frac{M}{N} \frac{1}{M} \sum_{j=0}^{N^*} k\left(\frac{j}{M}\right) \left(1 - \frac{M}{N}\frac{c}{M}\right)^j \\
&\xrightarrow{p} \sigma_{21} d_f \int_0^\infty k(r) e^{-cd_f r} dr.
\end{aligned}$$

When $M/N \to 0$, we can see that

$$\left| \frac{1}{N}\sum_{j=0}^{N^*} k\left(\frac{j}{M}\right) \frac{\rho^j}{T} \sum_{t=1}^{T-j} u_{2t}\varepsilon_{1t} \right| \le \frac{M}{N}\frac{1}{M} \sum_{j=0}^{N^*} \left| k\left(\frac{j}{M}\right) \right| \times O_p(1) \xrightarrow{p} 0.$$

(d) Since

$$\frac{1}{N}\hat{\omega}_{21} = \frac{1}{N}\sum_{j=1}^{T-1} k\left(\frac{j}{M}\right) \frac{1}{T} \sum_{t=1}^{T-j} u_{2t+j}\hat{u}_{1t} + \hat{\lambda}_{21}, \tag{40}$$

it is sufficient to show that the first term on the right hand side of (40) converges in probability to zero. Here note that

$$\frac{1}{N}\sum_{j=1}^{T-1}k\left(\frac{j}{M}\right)\frac{1}{T}\sum_{t=1}^{T-j}u_{2t+j}\hat{u}_{1t}=\frac{1}{N}\sum_{j=1}^{T-1}k\left(\frac{j}{M}\right)\frac{1}{T}\sum_{t=1}^{T-j}u_{2t+j}u_{1t}+O_p\left(\frac{M}{N}\right) \tag{41}$$

using (37). Since $E(u_{2t+j}u_{1t})=0$ for $j>0$ and

$$V\left(\sum_{t=1}^{T-j}u_{2t+j}u_{1t}\right)=\sum_{t=1}^{T-j}E(u_{1t}^2)E(u_{2t+j}^2)\leq C_4\times NT,$$

we note that the first term on the right-hand side of (41) is $O_p(K/(NT)^{1/2})$. Therefore, the first term on the right hand side of (40) converges in probability to zero.□

In order to derive the limiting distribution of the FMR estimator, we note that

$$\begin{aligned}
&\frac{1}{N}D_T(\hat{\theta}_{FMR}-\theta)\\
&= \left(D_T^{-1}\sum_{t=1}^{T}z_tz_t'D_T^{-1}\right)^{-1}\left(\frac{1}{N}D_T^{-1}\sum_{t=1}^{T}z_tu_{1t}-D_T^{-1}\sum_{t=1}^{T}z_tu_{2t}'\hat{\Omega}_{22}^{-1}\frac{1}{N}\hat{\omega}_{21}-\frac{1}{N}\hat{J}^+\right)\\
&= H_T^{-1}h_{FMR,T}, \quad \text{say.}
\end{aligned}$$

As we noted in the previous proofs, $H_T\Rightarrow H$ while noting the relation of $\Omega_{22}=\Lambda_{22}=\Sigma_{22}$ in our model, the limiting distribution of $h_{FMR,T}$ is derived using Lemmas A.1 and A.2 as follows:

$$\begin{aligned}
h_{FMR,T} &= \begin{bmatrix}\frac{1}{N\sqrt{T}}\sum_{t=1}^{T}u_{1t}-\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{2t}'\hat{\Omega}_{22}^{-1}\frac{1}{N}\hat{\omega}_{21}\\ \frac{1}{NT}\sum_{t=1}^{T}x_tu_{1t}-\frac{1}{T}\sum_{t=1}^{T}x_tu_{2t}'\hat{\Omega}_{22}^{-1}\frac{1}{N}\hat{\omega}_{21}-\frac{1}{N}\left(\hat{\lambda}_{21}-\hat{\Lambda}_{22}\hat{\Omega}_{22}^{-1}\hat{\omega}_{21}\right)\end{bmatrix}\\
&\Rightarrow \begin{bmatrix}\frac{1}{c}B_1(1)-B_2'(1)\Sigma_{22}^{-1}\bar{\lambda}_{21}\\ \frac{1}{c}\left(\int_0^1B_2(r)dB_1(r)+\sigma_{21}\right)-\left(\int_0^1B_2(r)dB_2'(r)+\Sigma_{22}\right)\Sigma_{22}^{-1}\bar{\lambda}_{21}\end{bmatrix}.
\end{aligned}$$

As in the proof of Theorem 1, the CCR estimator can be expressed as

$$\frac{1}{N}D_T(\hat{\theta}_{CCR}-\theta)=\begin{bmatrix}1 & \frac{1}{T\sqrt{T}}\sum_{t=1}^{T}x_t^{*\prime}\\ \frac{1}{T\sqrt{T}}\sum_{t=1}^{T}x_t^* & \frac{1}{T^2}\sum_{t=1}^{T}x_t^*x_t^{*\prime}\end{bmatrix}^{-1}\begin{bmatrix}\frac{1}{N\sqrt{T}}\sum_{t=1}^{T}u_{1t}^*\\ \frac{1}{NT}\sum_{t=1}^{T}x_t^*u_{1t}^*\end{bmatrix}.$$

From Lemma A.2, we obtain

$$\begin{aligned}
\hat{\Lambda}_2\hat{\Sigma}_u^{-1} &= \left[\frac{1}{N}\hat{\lambda}_{21},\hat{\Lambda}_{22}\right]\begin{bmatrix}\frac{1}{N}\hat{\sigma}_{11} & \hat{\sigma}_{12}\\ \frac{1}{N}\hat{\sigma}_{21} & \hat{\Sigma}_{22}\end{bmatrix}^{-1}\\
&\xrightarrow{p} \left[\bar{\lambda}_{21}\frac{2c}{\sigma_{11}},-\bar{\lambda}_{21}\frac{2c}{\sigma_{11}}\sigma_{12}\Sigma_{22}^{-1}+I_n\right].
\end{aligned} \tag{42}$$

Using (42), we can see that

$$\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}x_t^* = \frac{1}{T\sqrt{T}}\sum_{t=1}^{T}x_t - \hat{\Lambda}_2\hat{\Sigma}_u^{-1}\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}\hat{u}_t \Rightarrow \int_0^1 B_2(r)dr, \tag{43}$$

$$\frac{1}{T^2}\sum_{t=1}^{T}x_t^*x_t^{*\prime} \Rightarrow \int_0^1 B_2(r)B_2'(r)dr. \tag{44}$$

Further, using Lemmas A.1 and A.2, it is shown that

$$\begin{aligned}
&\frac{1}{N\sqrt{T}}\sum_{t=1}^{T}u_{1t}^* \\
&= \frac{1}{N\sqrt{T}}\sum_{t=1}^{T}u_{1t} - (\hat{\beta}-\beta)'\hat{\Lambda}_2\hat{\Sigma}_u^{-1}\frac{1}{N\sqrt{T}}\sum_{t=1}^{T}\hat{u}_t - \frac{1}{N}\hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{2t} \\
&\Rightarrow \frac{1}{c}B_1(1) - \bar{\lambda}_{12}\Sigma_{22}^{-1}B_2(1)
\end{aligned} \tag{45}$$

and

$$\begin{aligned}
\frac{1}{NT}\sum_{t=1}^{T}x_t^*u_{1t}^* &= \frac{1}{NT}\sum_{t=1}^{T}x_tu_{1t} - \frac{1}{NT}\sum_{t=1}^{T}x_t\hat{u}_t'\hat{\Sigma}_u^{-1}\hat{\Lambda}_2'(\hat{\beta}-\beta) \\
&\quad -\frac{1}{T}\sum_{t=1}^{T}x_tu_{2t}'\hat{\Omega}_{22}^{-1}\frac{1}{N}\hat{\omega}_{21} - \hat{\Lambda}_2\hat{\Sigma}_u^{-1}\frac{1}{NT}\sum_{t=1}^{T}\hat{u}_tu_{1t} \\
&\quad +\hat{\Lambda}_2\hat{\Sigma}_u^{-1}\frac{1}{NT}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t'\hat{\Sigma}_u^{-1}\hat{\Lambda}_2'(\hat{\beta}-\beta) + \hat{\Lambda}_2\hat{\Sigma}_u^{-1}\frac{1}{T}\sum_{t=1}^{T}\hat{u}_tu_{2t}'\hat{\Omega}_{22}^{-1}\frac{1}{N}\hat{\omega}_{21} \\
&\Rightarrow \frac{1}{c}\left(\int_0^1 B_2(r)dB_1(r) + \sigma_{21}\right) - \left(\int_0^1 B_2(r)dB_2'(r) + \Sigma_{22}\right)\Sigma_{22}^{-1}\bar{\lambda}_{21} \\
&\quad - \left[\bar{\lambda}_{21}\frac{2c}{\sigma_{11}}, -\bar{\lambda}_{21}\frac{2c}{\sigma_{11}}\sigma_{12}\Sigma_{22}^{-1} + I_n\right]\begin{bmatrix} \frac{\sigma_{11}}{2c} \\ 0 \end{bmatrix} \\
&\quad + \left[\bar{\lambda}_{21}\frac{2c}{\sigma_{11}}, -\bar{\lambda}_{21}\frac{2c}{\sigma_{11}}\sigma_{12}\Sigma_{22}^{-1} + I_n\right]\begin{bmatrix} \sigma_{12} \\ \Sigma_{22} \end{bmatrix}\Sigma_{22}^{-1}\bar{\lambda}_{21} \\
&= \left(\frac{1}{c}\int_0^1 B_2(r)dB_1(r) - \int_0^1 B_2(r)dB_2'(r)\Sigma_{22}^{-1}\bar{\lambda}_{21}\right) + \left(\frac{\sigma_{21}}{c} - \bar{\lambda}_{21}\right).
\end{aligned} \tag{46}$$

The limiting distribution is obtained using (43)–(46).

For the DOLS estimator, we first transform the error term ε_{1t} as

$$\varepsilon_{1t} = \varepsilon_{1\cdot 2t} + \tilde{\varepsilon}_{2t}$$

where $\varepsilon_{1\cdot 2t} = \varepsilon_{1t} - \sigma_{12}\Sigma_{22}^{-1}\varepsilon_{2t}$ and $\tilde{\varepsilon}_{2t} = \sigma_{12}\Sigma_{22}^{-1}\varepsilon_{2t}$. Notice that $\varepsilon_{1\cdot 2t}$ is uncorrelated with all the leads and lags of ε_{2t} and $\tilde{\varepsilon}_{2t}$. Since $u_{1t} = \sum_{l=1}^{t} \rho^{t-l}\varepsilon_{1l}$, we have

$$
\begin{aligned}
u_{1t} &= \sum_{l=1}^{t} \rho^{t-l}\varepsilon_{1\cdot 2l} + \sum_{l=1}^{t} \rho^{t-l}\tilde{\varepsilon}_{2l} \\
&= \sum_{l=t-K}^{t} \rho^{t-l}\tilde{\varepsilon}_{2l} + \sum_{l=1}^{t} \rho^{t-l}\varepsilon_{1\cdot 2l} + \sum_{l=1}^{t-K-1} \rho^{t-l}\tilde{\varepsilon}_{2l} \\
&= \sum_{j=0}^{K} \pi_j' \Delta x_{t-j} + u_{1\cdot 2t} + r_{2t} \\
&= \sum_{j=0}^{K} \pi_j' \Delta x_{t-j} + \dot{u}_{1\cdot 2t} \qquad (47)
\end{aligned}
$$

where $\pi_j' = \rho^j \sigma_{12}\Sigma_{22}^{-1}$, $u_{1\cdot 2t} = \sum_{l=1}^{t} \rho^{t-l}\varepsilon_{1\cdot 2l}$, $r_{2t} = \sum_{l=1}^{t-K-1} \rho^{t-l}\tilde{\varepsilon}_{2l}$, and $\dot{u}_{1\cdot 2t} = u_{1\cdot 2t} + r_{2t}$. Inserting (47) in the model, we obtain

$$
y_t = \theta' z_t + \sum_{j=0}^{K} \pi_j \Delta x_{t-j} + \dot{u}_{1\cdot 2t}.
$$

Then, the DOLS estimator can be expressed as

$$
\begin{aligned}
&\frac{1}{N} D_T(\hat{\theta}_{DOLS} - \theta) \\
&= \left(D_T^{-1} \sum_{t=K+1}^{T-K} z_t z_t' D_T^{-1} - G_1 G_2^{-1} G_1' \right)^{-1} \left(\frac{1}{N} D_T^{-1} \sum_{t=K+1}^{T-K} z_t \dot{u}_{1\cdot 2t} - G_1 G_2^{-1} G_4 \right) \qquad (48)
\end{aligned}
$$

where G_1 and G_2 are defined in the proof of Theorem 1 while

$$
G_4 = \frac{1}{N\sqrt{T}} \sum_{t=K+1}^{T-K} w_t \dot{u}_{1\cdot 2t}.
$$

The term in the first parentheses of (48) weakly converges to H^{-1} as proved in Theorem 1. In order to derive the limiting distribution of the term in the second parentheses of (48), we first investigate the order of $\|G_4\|$. Note that

$$
\begin{aligned}
\|G_4\|^2 &= \sum_{j=-K}^{K} \left\| \frac{1}{N\sqrt{T}} \sum_{t=K+1}^{T-K} \varepsilon_{2t-j} \dot{u}_{1\cdot 2t} \right\|^2 \\
&\leq 2 \sum_{j=-K}^{K} \left(\left\| \frac{1}{N\sqrt{T}} \sum_{t=K+1}^{T-K} \varepsilon_{2t-j} u_{1\cdot 2t} \right\|^2 + \left\| \frac{1}{N\sqrt{T}} \sum_{t=K+1}^{T-K} \varepsilon_{2t-j} r_{2t} \right\|^2 \right). \qquad (49)
\end{aligned}
$$

Here, we can see that $E(\varepsilon_{2t-j}u_{1\cdot 2t}) = 0$ and

$$\begin{aligned} Var\left(\sum_{t=K+1}^{T-K}\varepsilon_{2t-j}u_{1\cdot 2t}\right) &= \Sigma_{22}\sum_{t=K+1}^{T-K}E\left(\left(\sum_{l=1}^{t}\rho^{t-l}\varepsilon_{1\cdot 2l}\right)^2\right) \\ &= \Sigma_{22}\sigma_{1\cdot 2}\sum_{t=K+1}^{T-K}\frac{1-\rho^{2t}}{1-\rho^2} = O(NT) \end{aligned}$$

where the first equality holds because $E(\varepsilon_s\varepsilon_{1\cdot 2t}) = 0$ for all s and t and $\varepsilon_{1\cdot 2t}$ is an uncorrelated sequence. Then, the first term in the parentheses of (49) is $O_p(1/N)$. Since we can similarly show that the second term in the parentheses of (49) is $O_p(1/N)$, we can see that $\|G_4\|^2 = O_p(K/N)$. Since $\|G_1\|^2 = O_p(K/T)$ as in (29) and $\|G_2^{-1}\|_1 = O_p(1)$ as shown by Saikkonen (1991), we can see that

$$\|G_1G_2^{-1}G_4\| \leq O_p\left(\frac{\sqrt{K}}{\sqrt{T}}\right)O_p(1)\,O_p\left(\frac{\sqrt{K}}{\sqrt{N}}\right) = O_p\left(\frac{K}{\sqrt{NT}}\right) = o_p(1).$$

Then, the theorem is proved if we show that

$$\frac{1}{N}D_T^{-1}\sum_{t=K+1}^{T-K}z_t\dot{u}_{1\cdot 2t} = \frac{1}{N}D_T^{-1}\sum_{t=K+1}^{T-K}z_tu_{1\cdot 2t} + \frac{1}{N}D_T^{-1}\sum_{t=K+1}^{T-K}z_tr_{2t} \Rightarrow h_{DOLS}. \tag{50}$$

First, we derive the limiting distribution of the first term in the middle of (50). Note that

$$\begin{aligned} \sum_{t=K+1}^{T-K}u_{1\cdot 2t} &= \sum_{t=K+1}^{T-K}\sum_{l=1}^{t}\rho^{t-l}\varepsilon_{1\cdot 2l} \\ &= \sum_{j=1}^{K}\left(\rho^{K-j+1}+\rho^{K-j+2}+\cdots+\rho^{T-K-j}\right)\varepsilon_{1\cdot 2j} \\ &\quad + \sum_{j=1}^{T-2K}\left(1+\rho+\rho^2+\cdots+\rho^{T-2K-j}\right)\varepsilon_{1\cdot 2,K+j} \\ &= \sum_{j=1}^{K}\frac{\rho^{K-j+1}-\rho^{T-K-j+1}}{1-\rho}\varepsilon_{1\cdot 2j} \\ &\quad + \frac{1}{1-\rho}\sum_{j=1}^{T-2K}\varepsilon_{1\cdot 2,K+j} - \frac{1}{1-\rho}\sum_{j=1}^{T-2K}\rho^{T-2K-j+1}\varepsilon_{1\cdot 2,K+j}. \end{aligned} \tag{51}$$

From some algebra, we can show that the expectation of the first term of (51) is zero and its variance is $O(N^3)$ while those of the third term of (51) are zero and $O(N^3)$. Thus, we obtain

$$\frac{1}{N\sqrt{T}}\sum_{t=K+1}^{T-K} u_{1\cdot 2t} = \frac{1}{N(1-\rho)\sqrt{T}}\sum_{j=1}^{T-2K} \varepsilon_{1\cdot 2,K+j} + O_p\left(\sqrt{\frac{N}{T}}\right) \Rightarrow \frac{1}{c}B_{1\cdot 2}(1). \tag{52}$$

Similarly, we can show that

$$\frac{1}{NT}\sum_{t=K+1}^{T-K} x_t u_{1\cdot 2t} \Rightarrow \frac{1}{c}\int_0^1 B_2(r)dB_{1\cdot 2}(r). \tag{53}$$

For the second term in the middle of (50), we note that

$$\begin{aligned}
\sum_{t=K+1}^{T-K} r_{2t} &= \sum_{t=K+1}^{T-K}\sum_{l=1}^{t-K-1} \rho^{t-l}\tilde{\varepsilon}_{2l} \\
&= \sum_{j=1}^{T-2K-1}\left(\rho^{K+1}+\rho^{K+2}+\cdots+\rho^{T-K-j}\right)\tilde{\varepsilon}_{2j} \\
&= \frac{\rho^{K+1}}{1-\rho}\sum_{j=1}^{T-2K-1}\tilde{\varepsilon}_{2j} - \frac{1}{1-\rho}\sum_{j=1}^{T-2K-1}\rho^{T-K-j+1}\tilde{\varepsilon}_{2j}.
\end{aligned} \tag{54}$$

The second term of (54) is shown to be $O_p(N^{3/2})$ because its mean is zero and variance is $O(N^3)$. Thus, we obtain

$$\begin{aligned}
\frac{1}{N\sqrt{T}}\sum_{t=K+1}^{T-K} r_{2t} &= \frac{\rho^{K+1}}{N(1-\rho)\sqrt{T}}\sum_{j=1}^{T-2K-1}\tilde{\varepsilon}_{2j} + O_p\left(\sqrt{\frac{N}{T}}\right) \\
&\Rightarrow \begin{cases} 0 & : \frac{K}{N}\to\infty \\ \frac{1}{c}e^{-cd_d}\sigma_{12}\Sigma_{22}^{-1}B_2(1) & : \frac{K}{N}\to d_d \\ \frac{1}{c}\sigma_{12}\Sigma_{22}^{-1}B_2(1) & : \frac{K}{N}\to 0. \end{cases}
\end{aligned} \tag{55}$$

We can similarly show that

$$\frac{1}{NT}\sum_{t=K+1}^{T-K} z_t r_{2t} \Rightarrow \begin{cases} 0 & : \frac{K}{N}\to\infty \\ \frac{1}{c}e^{-cd_d}\left(\int_0^1 B_2(r)dB_2'(r)\Sigma_{22}^{-1}\sigma_{21}+\sigma_{21}\right) & : \frac{K}{N}\to d_d \\ \frac{1}{c}\left(\int_0^1 B_2(r)B_2'(r)\Sigma_{22}^{-1}\sigma_{21}+\sigma_{21}\right) & : \frac{K}{N}\to 0. \end{cases} \tag{56}$$

By combining (52), (53), (55), and (56), we obtain (50).■

References

[1] Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* 59, 817-858.

[2] Andrews, D. W. K. and J. C. Monahan (1992). An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator. *Econometrica* 60, 953-966.

[3] Brillinger, D. R. (1981). *Time Series Data Analysis and Theory*. Holden-Day, San Francisco.

[4] Cappuccio, N. and D. Lubian (2001). Estimation and Inference on Long-Run Equilibria: A Simulation Study. *Econometric Reviews* 20, 61-84.

[5] Christou, C. and N. Pittis (2002). Kernel and Bandwidth Selection, Prewhitening, and the Performance of the Fully Modified Least Squares Estimation Method. *Econometric Theory* 18, 948-961.

[6] Engle, R. F. and C. W. J. Granger (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* 55, 251-276.

[7] Giraitis, L. and P. C. B. Phillips (2006). Uniform Limit Theory for Stationary Autoregression. *Journal of Time Series Analysis* 27, 51-60.

[8] Hansen, B. E. (1992). Convergence to Stochastic Integrals for Dependent Heterogeneous Processes. *Econometric Theory* 8, 489-500.

[9] Inder, B. (1993). Estimating Long-Run Relationships in Economics. *Journal of Econometrics* 57, 53-68.

[10] Jansson, M. (2002). Consistent Covariance Matrix Estimation for Linear Processes. *Econometric Theory* 18, 1449-1459.

[11] Kitamura, Y. and P. C. B. Phillips (1997). Fully Modified IV, GIVE and GMM Estimation with Possibly Non-Stationary Regressors and Instruments. *Journal of Econometrics* 80, 85-123.

[12] Kejriwal, M. and P. Perron (2006). Data Dependent Rules for the Selection of the Number of Leads and Lags in the Dynamic OLS Cointegrating Regression," manuscript.

[13] Montalvo, J. G. (1995). Comparing Cointegrating Regression Estimators: Some Additional Monte Carlo Results. *Economics Letters* 48, 229-234.

[14] Newey, W. K. and K. D. West (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies* 61, 631-653.

[15] Ng, S. and P. Perron (1995). Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag. *Journal of the American Statistical Society* 90, 268-281.

[16] Park, J. Y. (1992) Canonical Cointegrating Regressions. *Econometrica* 60, 119-143.

[17] Phillips, P. C. B. (1987). Towards a Unified Asymptotic Theory for Autoregression. *Biometrika* 74, 535-547.

[18] Phillips, P. C. B. (1995). Fully Modified Least Squares and Vector Autoregression. *Econometrica* 63, 1023-1078.

[19] Phillips, P. C. B. and B. E. Hansen (1990). Statistical Inference in Instrumental Variables Regression With I(1) Processes. *Review of Economic Studies* 57, 99-125.

[20] Phillips, P. C. B. and M. Loretan (1991). Estimating Long-Run Economic Equilibria. *Review of Economic Studies* 58, 407-436.

[21] Phillips, P. C. B. and T. Magdalinos (2005). Limit Theory for Moderate Deviations from a Unit Root under Weak Dependence. Cowles Foundation Discussion paper No. 1517.

[22] Phillips, P. C. B. and T. Magdalinos (2007). Limit Theory for Moderate Deviations from a Unit Root. *Journal of Econometrics* 136, 115-130.

[23] Saikkonen, P. (1991) Asymptotically Efficient Estimation of Cointegration Regressions. *Econometric Theory* 7, 1-21.

[24] Saikkonen, P. and H. Lütkepohl (1999). Local Power of Likelihood Ratio Tests for the Cointegrating Rank of a VAR Process. *Econometric Theory* 15, 50-78.

[25] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

[26] Stock, J. H. and M. W. Watson (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica* 61, 783-820.

[27] Sul, D., P. C. B. Phillips and C. Y. Choi (2005). Prewhitening Bias in HAC Estimation. *Oxford Bulletin of Economics and Statistics* 67, 517-546.

[28] Tanaka, K. (1996). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. Wiley, New York.

Table 1a. Bias of the estimators ($T = 100$, $\sigma_{11} = 1$, $\sigma_{22} = 1$, $\sigma_{21} = 0.4$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.02328	0.00378	0.00362	0.00501	0.00420	0.00339	0.00339	0.00507	0.00419
0.3	0.03020	0.00880	0.00813	0.00935	0.00850	0.00556	0.00556	0.00753	0.00649
0.5	0.03960	0.01486	0.01352	0.01554	0.01461	0.00766	0.00768	0.00982	0.00874
0.7	0.06183	0.03064	0.02867	0.03165	0.03186	0.01477	0.01482	0.01736	0.01619
0.8	0.08584	0.05164	0.04981	0.05366	0.05548	0.02829	0.02845	0.03098	0.02936
0.85	0.10539	0.06962	0.06811	0.07355	0.07612	0.04045	0.04064	0.04331	0.04141
0.9	0.13920	0.10231	0.10139	0.10920	0.11268	0.06794	0.06824	0.07085	0.06896
0.95	0.20944	0.17658	0.17633	0.18663	0.19036	0.14672	0.14681	0.14842	0.14774
0.98	0.29221	0.26959	0.26903	0.27992	0.28243	0.25802	0.25801	0.25806	0.25726
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00332	0.00318	0.00454	0.00374	0.00301	0.00302	0.00463	0.00377
0.3		0.00839	0.00776	0.00895	0.00814	0.00546	0.00547	0.00729	0.00630
0.5		0.01460	0.01339	0.01528	0.01441	0.00826	0.00829	0.01016	0.00921
0.7		0.03078	0.02918	0.03162	0.03181	0.01783	0.01790	0.01984	0.01893
0.8		0.05228	0.05106	0.05378	0.05543	0.03427	0.03442	0.03642	0.03512
0.85		0.07071	0.07002	0.07369	0.07606	0.04952	0.04969	0.05166	0.05024
0.9		0.10391	0.10397	0.10936	0.11260	0.08206	0.08227	0.08391	0.08255
0.95		0.17891	0.17930	0.18680	0.19027	0.16747	0.16745	0.16799	0.16732
0.98		0.27164	0.27168	0.28004	0.28239	0.27636	0.27634	0.27620	0.27523
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00128	0.00060	0.00090	0.00223				
0.3		0.00637	0.00256	0.00379	0.00871				
0.5		0.01211	0.00516	0.00740	0.01721				
0.7		0.02131	0.00819	0.01181	0.03634				
0.8		0.03907	0.02225	0.02796	0.05952				
0.85		0.05457	0.03461	0.04166	0.07897				
0.9		0.08836	0.07096	0.07832	0.11442				
0.95		0.16391	0.15376	0.16105	0.18505				
0.98		0.26970	0.26051	0.26842	0.27621				

Table 1b. Bias of the estimators ($T = 100$, $\sigma_{11} = 1$, $\sigma_{22} = 1$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.04609	0.00755	0.00723	0.00981	0.00820	0.00664	0.00661	0.01044	0.00859
0.3	0.05769	0.01518	0.01380	0.01598	0.01421	0.00761	0.00759	0.01272	0.01021
0.5	0.07852	0.02939	0.02691	0.02989	0.02803	0.01116	0.01117	0.01830	0.01466
0.7	0.12321	0.06288	0.05982	0.06351	0.06391	0.02369	0.02377	0.03179	0.02750
0.8	0.16947	0.10277	0.10087	0.10507	0.10846	0.04068	0.04084	0.04992	0.04482
0.85	0.21324	0.14247	0.14207	0.14818	0.15367	0.06613	0.06647	0.07515	0.07019
0.9	0.28403	0.21313	0.21509	0.22359	0.23079	0.12160	0.12207	0.12974	0.12563
0.95	0.41875	0.35536	0.35903	0.37316	0.38048	0.27455	0.27490	0.27991	0.27689
0.98	0.59282	0.54900	0.54997	0.56950	0.57402	0.52499	0.52534	0.52570	0.52438
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00678	0.00659	0.00866	0.00729	0.00630	0.00627	0.00936	0.00773
0.3		0.01565	0.01475	0.01629	0.01495	0.01125	0.01127	0.01453	0.01268
0.5		0.03160	0.03009	0.03203	0.03065	0.02259	0.02268	0.02580	0.02387
0.7		0.06751	0.06593	0.06766	0.06777	0.05161	0.05182	0.05386	0.05245
0.8		0.10869	0.10839	0.10936	0.11170	0.08610	0.08639	0.08813	0.08663
0.85		0.14910	0.15002	0.15199	0.15613	0.12381	0.12415	0.12577	0.12430
0.9		0.22019	0.22322	0.22619	0.23203	0.19260	0.19297	0.19430	0.19324
0.95		0.36223	0.36653	0.37417	0.38067	0.34583	0.34597	0.34642	0.34549
0.98		0.55421	0.55601	0.56976	0.57396	0.56606	0.56610	0.56528	0.56534
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00274	0.00076	0.00181	0.00408				
0.3		0.00399	-0.00020	0.00158	0.00655				
0.5		0.00896	0.00319	0.00461	0.01371				
0.7		0.02193	0.01090	0.01125	0.03522				
0.8		0.04393	0.02786	0.02937	0.06546				
0.85		0.07413	0.05296	0.05470	0.10286				
0.9		0.15287	0.13049	0.13257	0.18452				
0.95		0.32049	0.30618	0.31121	0.35801				
0.98		0.54549	0.54061	0.54920	0.56944				

Table 1c. Bias of the estimators ($T = 100$, $\sigma_{11} = 1$, $\sigma_{22} = 3$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.01531	0.00233	0.00222	0.00329	0.00270	0.00209	0.00208	0.00325	0.00264
0.3	0.01925	0.00515	0.00474	0.00571	0.00505	0.00315	0.00316	0.00450	0.00375
0.5	0.02649	0.00985	0.00897	0.01065	0.01015	0.00495	0.00498	0.00660	0.00573
0.7	0.04008	0.01967	0.01837	0.02105	0.02120	0.00889	0.00893	0.01094	0.00971
0.8	0.05779	0.03487	0.03358	0.03714	0.03827	0.01859	0.01870	0.02112	0.01976
0.85	0.07141	0.04718	0.04608	0.05069	0.05247	0.02803	0.02816	0.03058	0.02885
0.9	0.09478	0.07027	0.06964	0.07538	0.07779	0.04735	0.04758	0.04988	0.04852
0.95	0.13840	0.11574	0.11566	0.12353	0.12593	0.09487	0.09518	0.09600	0.09504
0.98	0.19279	0.17744	0.17731	0.18471	0.18628	0.16674	0.16679	0.16763	0.16680
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00203	0.00192	0.00296	0.00239	0.00184	0.00184	0.00295	0.00235
0.3		0.00491	0.00453	0.00546	0.00483	0.00313	0.00314	0.00436	0.00366
0.5		0.00973	0.00896	0.01052	0.01006	0.00557	0.00561	0.00695	0.00620
0.7		0.01989	0.01885	0.02112	0.02124	0.01146	0.01151	0.01299	0.01207
0.8		0.03540	0.03454	0.03728	0.03829	0.02359	0.02370	0.02536	0.02437
0.85		0.04801	0.04746	0.05083	0.05246	0.03508	0.03521	0.03667	0.03555
0.9		0.07140	0.07136	0.07550	0.07774	0.05797	0.05814	0.05943	0.05863
0.95		0.11740	0.11787	0.12363	0.12587	0.11075	0.11088	0.11111	0.11075
0.98		0.17907	0.17943	0.18475	0.18624	0.17971	0.17980	0.17943	0.17955
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00068	0.00020	0.00040	0.00138				
0.3		0.00348	0.00146	0.00178	0.00514				
0.5		0.00631	0.00185	0.00284	0.01057				
0.7		0.01315	0.00507	0.00722	0.02183				
0.8		0.02442	0.01297	0.01558	0.03857				
0.85		0.03415	0.02190	0.02495	0.05092				
0.9		0.05626	0.04555	0.04865	0.07476				
0.95		0.10631	0.09797	0.10205	0.12149				
0.98		0.17567	0.17414	0.17794	0.18298				

Table 1d. Bias of the estimators ($T = 100$, $\sigma_{11} = 1$, $\sigma_{22} = 3$, $\sigma_{21} = 1.6$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.03061	0.00471	0.00450	0.00654	0.00533	0.00414	0.00412	0.00725	0.00582
0.3	0.03921	0.01044	0.00962	0.01133	0.01013	0.00438	0.00438	0.00941	0.00738
0.5	0.05214	0.01928	0.01769	0.02046	0.01945	0.00518	0.00526	0.01236	0.00917
0.7	0.08242	0.04240	0.04066	0.04387	0.04429	0.01077	0.01101	0.02047	0.01595
0.8	0.11486	0.07102	0.07067	0.07318	0.07560	0.01996	0.02034	0.03123	0.02596
0.85	0.14041	0.09486	0.09587	0.09888	0.10258	0.03200	0.03246	0.04334	0.03813
0.9	0.18731	0.14172	0.14509	0.14822	0.15295	0.06369	0.06419	0.07503	0.06968
0.95	0.27729	0.23709	0.24168	0.24839	0.25309	0.16917	0.16953	0.17576	0.17240
0.98	0.38757	0.35875	0.36101	0.37232	0.37518	0.33125	0.33136	0.33139	0.33054
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00432	0.00427	0.00544	0.00468	0.00439	0.00439	0.00617	0.00507
0.3		0.01201	0.01181	0.01239	0.01183	0.01093	0.01096	0.01257	0.01154
0.5		0.02269	0.02216	0.02317	0.02261	0.01985	0.01997	0.02136	0.02042
0.7		0.04759	0.04711	0.04805	0.04816	0.04154	0.04181	0.04280	0.04212
0.8		0.07703	0.07760	0.07751	0.07889	0.06807	0.06840	0.06929	0.06865
0.85		0.10120	0.10280	0.10269	0.10510	0.09105	0.09141	0.09239	0.09181
0.9		0.14800	0.15146	0.15089	0.15441	0.13584	0.13617	0.13733	0.13675
0.95		0.24249	0.24690	0.24931	0.25336	0.23635	0.23654	0.23709	0.23662
0.98		0.36279	0.36492	0.37253	0.37516	0.37106	0.37103	0.37004	0.37005
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00111	0.00001	0.00032	0.00176				
0.3		0.00190	0.00057	0.00089	0.00293				
0.5		0.00367	0.00138	0.00164	0.00544				
0.7		0.00937	0.00471	0.00484	0.01391				
0.8		0.02199	0.01617	0.01658	0.02856				
0.85		0.03899	0.03189	0.03263	0.04711				
0.9		0.08724	0.07907	0.07958	0.09774				
0.95		0.20813	0.19979	0.20082	0.22461				
0.98		0.35609	0.35039	0.35538	0.36815				

Table 1e. Bias of the estimators ($T = 100$, $\sigma_{11} = 3$, $\sigma_{22} = 1$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.04682	0.00758	0.00733	0.00995	0.00828	0.00675	0.00674	0.01010	0.00822
0.3	0.05730	0.01439	0.01294	0.01505	0.01339	0.00781	0.00780	0.01154	0.00950
0.5	0.08178	0.03146	0.02874	0.03203	0.02982	0.01652	0.01651	0.02053	0.01835
0.7	0.12376	0.06297	0.05933	0.06356	0.06373	0.03086	0.03090	0.03544	0.03286
0.8	0.16727	0.10004	0.09648	0.10323	0.10628	0.05227	0.05242	0.05669	0.05386
0.85	0.21018	0.13642	0.13264	0.14426	0.14984	0.07886	0.07913	0.08214	0.08017
0.9	0.28882	0.21350	0.21154	0.22710	0.23470	0.14700	0.14748	0.15058	0.14861
0.95	0.41828	0.35333	0.35374	0.37219	0.38005	0.29291	0.29321	0.29424	0.29290
0.98	0.58924	0.54352	0.54208	0.56509	0.56963	0.53380	0.53507	0.53443	0.53279
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00669	0.00648	0.00901	0.00737	0.00603	0.00602	0.00921	0.00739
0.3		0.01367	0.01234	0.01436	0.01282	0.00782	0.00783	0.01120	0.00931
0.5		0.03115	0.02876	0.03176	0.02972	0.01854	0.01856	0.02184	0.02001
0.7		0.06366	0.06082	0.06396	0.06407	0.03871	0.03882	0.04209	0.04016
0.8		0.10163	0.09939	0.10380	0.10654	0.06655	0.06673	0.06981	0.06773
0.85		0.13898	0.13692	0.14491	0.14997	0.10026	0.10058	0.10240	0.10113
0.9		0.21710	0.21689	0.22773	0.23472	0.17759	0.17799	0.17961	0.17852
0.95		0.35802	0.35987	0.37279	0.38008	0.33437	0.33473	0.33476	0.33430
0.98		0.54848	0.54838	0.56532	0.56951	0.55950	0.55993	0.55907	0.55925
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00305	0.00188	0.00246	0.00431				
0.3		0.00880	0.00036	0.00495	0.01383				
0.5		0.01984	0.00672	0.01276	0.03320				
0.7		0.04228	0.01496	0.02218	0.06964				
0.8		0.07085	0.03811	0.04579	0.10991				
0.85		0.10031	0.06619	0.07561	0.15171				
0.9		0.18408	0.14707	0.15801	0.22520				
0.95		0.32694	0.30571	0.30952	0.37409				
0.98		0.54998	0.54365	0.54641	0.56682				

Table 1f. Bias of the estimators ($T = 100$, $\sigma_{11} = 3$, $\sigma_{22} = 1$, $\sigma_{21} = 1.6$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.09239	0.01549	0.01487	0.01867	0.01595	0.01243	0.01231	0.02137	0.01732
0.3	0.11527	0.03107	0.02828	0.03139	0.02834	0.01221	0.01203	0.02567	0.02011
0.5	0.15707	0.05978	0.05494	0.05909	0.05494	0.01720	0.01699	0.03516	0.02672
0.7	0.24496	0.12713	0.12235	0.12501	0.12547	0.03535	0.03520	0.05762	0.04573
0.8	0.34307	0.21131	0.21002	0.21213	0.21879	0.06545	0.06533	0.08856	0.07583
0.85	0.42162	0.28587	0.28922	0.29075	0.30129	0.09929	0.09948	0.12145	0.10875
0.9	0.56273	0.42718	0.43751	0.43934	0.45442	0.20241	0.20288	0.22460	0.21197
0.95	0.83632	0.71402	0.72828	0.74354	0.75927	0.51981	0.52055	0.53596	0.52583
0.98	1.16945	1.08272	1.09054	1.11965	1.12931	1.02246	1.02277	1.02413	1.02121
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.01430	0.01418	0.01603	0.01447	0.01342	0.01339	0.01840	0.01546
0.3		0.03540	0.03454	0.03535	0.03401	0.03091	0.03102	0.03517	0.03242
0.5		0.06947	0.06778	0.06895	0.06666	0.06018	0.06044	0.06325	0.06092
0.7		0.14227	0.14109	0.13996	0.13975	0.12529	0.12580	0.12670	0.12508
0.8		0.22946	0.23103	0.22739	0.23099	0.20574	0.20635	0.20558	0.20482
0.85		0.30483	0.31012	0.30425	0.31076	0.27613	0.27680	0.27591	0.27488
0.9		0.44563	0.45616	0.44900	0.45981	0.41301	0.41366	0.41388	0.41230
0.95		0.73014	0.74364	0.74731	0.76052	0.71565	0.71614	0.71706	0.71487
0.98		1.09429	1.10193	1.12061	1.12935	1.12139	1.12145	1.11878	1.11860
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00390	0.00137	0.00185	0.00590				
0.3		0.00425	0.00093	0.00205	0.00778				
0.5		0.00995	0.00348	0.00458	0.01602				
0.7		0.02849	0.01506	0.01456	0.04160				
0.8		0.06338	0.04548	0.04589	0.08378				
0.85		0.11634	0.09526	0.09520	0.14063				
0.9		0.25742	0.23400	0.23631	0.29273				
0.95		0.63491	0.60607	0.61157	0.67960				
0.98		1.08236	1.06816	1.08338	1.11893				

Table 1g. Bias of the estimators ($T = 300$, $\sigma_{11} = 1$, $\sigma_{22} = 1$, $\sigma_{21} = 0.4$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00781	0.00077	0.00070	0.00102	0.00078	0.00049	0.00049	0.00091	0.00070
0.3	0.01031	0.00180	0.00152	0.00203	0.00172	0.00081	0.00081	0.00131	0.00107
0.5	0.01350	0.00281	0.00220	0.00318	0.00266	0.00059	0.00060	0.00119	0.00088
0.7	0.02276	0.00755	0.00636	0.00776	0.00779	0.00238	0.00239	0.00327	0.00278
0.8	0.03348	0.01412	0.01243	0.01485	0.01596	0.00476	0.00478	0.00598	0.00532
0.85	0.04255	0.01935	0.01730	0.02102	0.02346	0.00622	0.00625	0.00755	0.00693
0.9	0.06164	0.03346	0.03109	0.03761	0.04186	0.01347	0.01352	0.01471	0.01412
0.95	0.10254	0.06817	0.06605	0.07834	0.08440	0.03482	0.03489	0.03664	0.03550
0.98	0.18743	0.15165	0.14891	0.16917	0.17477	0.12218	0.12228	0.12201	0.12181
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00072	0.00065	0.00097	0.00073	0.00044	0.00044	0.00087	0.00066
0.3		0.00176	0.00149	0.00199	0.00169	0.00080	0.00080	0.00129	0.00105
0.5		0.00282	0.00223	0.00318	0.00268	0.00070	0.00071	0.00126	0.00097
0.7		0.00771	0.00660	0.00790	0.00791	0.00294	0.00295	0.00374	0.00330
0.8		0.01444	0.01293	0.01509	0.01612	0.00613	0.00615	0.00716	0.00660
0.85		0.01992	0.01816	0.02135	0.02363	0.00868	0.00871	0.00975	0.00925
0.9		0.03447	0.03263	0.03800	0.04200	0.01807	0.01812	0.01904	0.01862
0.95		0.07008	0.06888	0.07867	0.08446	0.04577	0.04584	0.04706	0.04621
0.98		0.15475	0.15366	0.16939	0.17478	0.13807	0.13813	0.13784	0.13769
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00067	0.00035	0.00038	0.00082				
0.3		0.00198	0.00087	0.00111	0.00274				
0.5		0.00289	0.00098	0.00129	0.00444				
0.7		0.00787	0.00404	0.00387	0.01135				
0.8		0.01549	0.00907	0.00888	0.02134				
0.85		0.02210	0.01308	0.01257	0.02995				
0.9		0.03817	0.02752	0.02692	0.04935				
0.95		0.08072	0.06804	0.06798	0.09235				
0.98		0.17121	0.16146	0.16184	0.18112				

Table 1h. Bias of the estimators ($T = 300$, $\sigma_{11} = 1$, $\sigma_{22} = 1$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.01575	0.00134	0.00120	0.00181	0.00138	0.00075	0.00075	0.00161	0.00122
0.3	0.02027	0.00315	0.00255	0.00356	0.00297	0.00093	0.00094	0.00209	0.00157
0.5	0.02818	0.00655	0.00532	0.00724	0.00619	0.00150	0.00150	0.00316	0.00242
0.7	0.04551	0.01524	0.01292	0.01561	0.01564	0.00315	0.00316	0.00565	0.00440
0.8	0.06618	0.02679	0.02350	0.02819	0.03068	0.00526	0.00528	0.00832	0.00685
0.85	0.08669	0.04064	0.03689	0.04358	0.04830	0.00839	0.00843	0.01239	0.01040
0.9	0.12184	0.06685	0.06295	0.07419	0.08236	0.01734	0.01741	0.02205	0.01990
0.95	0.20880	0.14207	0.14043	0.16048	0.17264	0.05709	0.05720	0.06285	0.06031
0.98	0.37399	0.30565	0.30114	0.33840	0.34920	0.23868	0.23886	0.24060	0.23928
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00130	0.00118	0.00172	0.00134	0.00079	0.00079	0.00155	0.00120
0.3		0.00344	0.00294	0.00381	0.00329	0.00165	0.00166	0.00256	0.00213
0.5		0.00742	0.00644	0.00803	0.00716	0.00375	0.00376	0.00482	0.00430
0.7		0.01725	0.01556	0.01754	0.01755	0.00972	0.00974	0.01087	0.01025
0.8		0.03016	0.02793	0.03112	0.03298	0.01849	0.01853	0.01952	0.01900
0.85		0.04506	0.04265	0.04701	0.05059	0.02938	0.02943	0.03038	0.02982
0.9		0.07279	0.07053	0.07769	0.08419	0.05190	0.05199	0.05285	0.05235
0.95		0.15015	0.15013	0.16293	0.17348	0.11986	0.11999	0.12079	0.12021
0.98		0.31437	0.31306	0.33928	0.34937	0.28809	0.28822	0.28881	0.28833
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00078	0.00020	0.00030	0.00114				
0.3		0.00111	0.00038	0.00051	0.00156				
0.5		0.00236	0.00106	0.00122	0.00326				
0.7		0.00562	0.00305	0.00300	0.00800				
0.8		0.01060	0.00662	0.00596	0.01517				
0.85		0.01844	0.01339	0.01248	0.02489				
0.9		0.03959	0.03307	0.03190	0.04867				
0.95		0.12568	0.11592	0.11428	0.14471				
0.98		0.33288	0.31936	0.31804	0.35121				

Table 1i. Bias of the estimators ($T = 300$, $\sigma_{11} = 1$, $\sigma_{22} = 3$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00522	0.00043	0.00040	0.00059	0.00045	0.00024	0.00025	0.00051	0.00039
0.3	0.00667	0.00104	0.00086	0.00118	0.00097	0.00034	0.00035	0.00069	0.00052
0.5	0.00921	0.00210	0.00170	0.00245	0.00218	0.00062	0.00062	0.00107	0.00086
0.7	0.01547	0.00530	0.00450	0.00569	0.00573	0.00175	0.00176	0.00238	0.00203
0.8	0.02181	0.00876	0.00763	0.00947	0.01031	0.00250	0.00252	0.00334	0.00284
0.85	0.02834	0.01297	0.01162	0.01426	0.01590	0.00394	0.00397	0.00485	0.00430
0.9	0.03969	0.02128	0.01977	0.02405	0.02682	0.00764	0.00769	0.00894	0.00826
0.95	0.06901	0.04613	0.04474	0.05306	0.05706	0.02355	0.02360	0.02474	0.02406
0.98	0.12180	0.09822	0.09622	0.10988	0.11354	0.07782	0.07785	0.07814	0.07770
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00041	0.00037	0.00056	0.00042	0.00022	0.00022	0.00048	0.00036
0.3		0.00102	0.00085	0.00116	0.00095	0.00035	0.00036	0.00069	0.00052
0.5		0.00212	0.00175	0.00247	0.00221	0.00073	0.00074	0.00114	0.00096
0.7		0.00544	0.00471	0.00581	0.00584	0.00226	0.00227	0.00279	0.00249
0.8		0.00908	0.00810	0.00970	0.01046	0.00374	0.00376	0.00440	0.00400
0.85		0.01346	0.01233	0.01456	0.01607	0.00602	0.00604	0.00670	0.00628
0.9		0.02206	0.02091	0.02438	0.02695	0.01147	0.01152	0.01241	0.01192
0.95		0.04756	0.04681	0.05332	0.05712	0.03186	0.03191	0.03268	0.03220
0.98		0.10035	0.09950	0.11001	0.11355	0.08872	0.08875	0.08887	0.08865
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00037	0.00015	0.00021	0.00048				
0.3		0.00104	0.00039	0.00040	0.00144				
0.5		0.00182	0.00064	0.00088	0.00267				
0.7		0.00466	0.00233	0.00247	0.00677				
0.8		0.00802	0.00425	0.00415	0.01180				
0.85		0.01275	0.00779	0.00754	0.01808				
0.9		0.02216	0.01529	0.01487	0.02989				
0.95		0.05319	0.04505	0.04436	0.06141				
0.98		0.11081	0.10493	0.10535	0.11656				

Table 1j. Bias of the estimators ($T = 300$, $\sigma_{11} = 1$, $\sigma_{22} = 3$, $\sigma_{21} = 1.6$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.01047	0.00089	0.00082	0.00121	0.00090	0.00044	0.00044	0.00109	0.00079
0.3	0.01331	0.00211	0.00175	0.00237	0.00198	0.00039	0.00039	0.00139	0.00100
0.5	0.01870	0.00431	0.00351	0.00494	0.00436	0.00052	0.00054	0.00200	0.00138
0.7	0.02997	0.00978	0.00824	0.01035	0.01045	0.00082	0.00086	0.00312	0.00214
0.8	0.04396	0.01799	0.01588	0.01914	0.02087	0.00167	0.00173	0.00479	0.00338
0.85	0.05675	0.02657	0.02422	0.02882	0.03210	0.00280	0.00290	0.00669	0.00498
0.9	0.08083	0.04479	0.04263	0.04972	0.05523	0.00678	0.00691	0.01167	0.00952
0.95	0.13963	0.09637	0.09661	0.10790	0.11583	0.02731	0.02749	0.03387	0.03106
0.98	0.25197	0.20761	0.20599	0.22892	0.23590	0.15742	0.15758	0.15971	0.15844
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00094	0.00090	0.00119	0.00096	0.00067	0.00067	0.00112	0.00089
0.3		0.00263	0.00240	0.00281	0.00256	0.00166	0.00167	0.00220	0.00196
0.5		0.00543	0.00493	0.00589	0.00549	0.00343	0.00345	0.00407	0.00377
0.7		0.01192	0.01100	0.01231	0.01235	0.00786	0.00790	0.00851	0.00823
0.8		0.02124	0.02004	0.02191	0.02303	0.01497	0.01502	0.01555	0.01529
0.85		0.03067	0.02942	0.03198	0.03420	0.02287	0.02294	0.02338	0.02320
0.9		0.05012	0.04914	0.05297	0.05696	0.04002	0.04010	0.04039	0.04029
0.95		0.10332	0.10427	0.11027	0.11672	0.08902	0.08912	0.08931	0.08923
0.98		0.21494	0.21526	0.22970	0.23607	0.20051	0.20060	0.20118	0.20083
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00010	-0.00001	0.00004	0.00023				
0.3		0.00040	0.00018	0.00017	0.00062				
0.5		0.00090	0.00045	0.00047	0.00126				
0.7		0.00196	0.00103	0.00092	0.00281				
0.8		0.00442	0.00305	0.00277	0.00588				
0.85		0.00782	0.00644	0.00612	0.00965				
0.9		0.02059	0.01954	0.01931	0.02244				
0.95		0.07573	0.07419	0.07399	0.07901				
0.98		0.21837	0.21215	0.21092	0.22865				

Table 1k. Bias of the estimators ($T = 300$, $\sigma_{11} = 3$, $\sigma_{22} = 1$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.01594	0.00159	0.00146	0.00206	0.00161	0.00102	0.00102	0.00186	0.00145
0.3	0.01963	0.00295	0.00237	0.00337	0.00273	0.00094	0.00094	0.00194	0.00147
0.5	0.02835	0.00661	0.00537	0.00709	0.00602	0.00200	0.00201	0.00320	0.00257
0.7	0.04550	0.01474	0.01232	0.01507	0.01511	0.00422	0.00423	0.00577	0.00494
0.8	0.06525	0.02577	0.02238	0.02702	0.02937	0.00705	0.00707	0.00901	0.00794
0.85	0.08735	0.04083	0.03674	0.04377	0.04845	0.01383	0.01388	0.01607	0.01495
0.9	0.12165	0.06498	0.06011	0.07331	0.08153	0.02458	0.02462	0.02721	0.02606
0.95	0.21375	0.14483	0.14071	0.16474	0.17720	0.07797	0.07809	0.08021	0.07910
0.98	0.38434	0.31597	0.30977	0.35032	0.36069	0.27075	0.27092	0.27145	0.27111
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00151	0.00137	0.00196	0.00152	0.00095	0.00095	0.00176	0.00136
0.3		0.00290	0.00234	0.00331	0.00269	0.00096	0.00096	0.00193	0.00147
0.5		0.00668	0.00551	0.00717	0.00613	0.00236	0.00236	0.00346	0.00287
0.7		0.01520	0.01301	0.01549	0.01551	0.00581	0.00583	0.00712	0.00642
0.8		0.02675	0.02381	0.02781	0.02994	0.01090	0.01093	0.01245	0.01162
0.85		0.04230	0.03889	0.04473	0.04901	0.02016	0.02023	0.02182	0.02101
0.9		0.06744	0.06371	0.07437	0.08197	0.03620	0.03625	0.03809	0.03726
0.95		0.14926	0.14713	0.16561	0.17739	0.10297	0.10311	0.10445	0.10361
0.98		0.32217	0.31912	0.35075	0.36071	0.29793	0.29805	0.29860	0.29826
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00139	0.00065	0.00092	0.00173				
0.3		0.00289	0.00094	0.00120	0.00425				
0.5		0.00561	0.00222	0.00271	0.00827				
0.7		0.01279	0.00569	0.00581	0.01897				
0.8		0.02388	0.01248	0.01233	0.03532				
0.85		0.04068	0.02561	0.02462	0.05649				
0.9		0.06886	0.04854	0.04575	0.09188				
0.95		0.16384	0.13971	0.13892	0.19006				
0.98		0.35496	0.33767	0.34133	0.37042				

Table 11. Bias of the estimators ($T = 300$, $\sigma_{11} = 3$, $\sigma_{22} = 1$, $\sigma_{21} = 1.6$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.03144	0.00266	0.00237	0.00347	0.00266	0.00127	0.00127	0.00321	0.00238
0.3	0.03996	0.00612	0.00491	0.00684	0.00562	0.00104	0.00103	0.00401	0.00285
0.5	0.05582	0.01294	0.01049	0.01377	0.01159	0.00169	0.00165	0.00599	0.00416
0.7	0.09080	0.02977	0.02507	0.02999	0.02994	0.00294	0.00289	0.00891	0.00609
0.8	0.13151	0.05385	0.04750	0.05576	0.06035	0.00579	0.00573	0.01395	0.00992
0.85	0.17069	0.07869	0.07150	0.08393	0.09341	0.00829	0.00824	0.01696	0.01289
0.9	0.24168	0.13422	0.12797	0.14659	0.16272	0.02140	0.02132	0.03288	0.02759
0.95	0.42274	0.29254	0.29367	0.32522	0.34981	0.09150	0.09154	0.10624	0.09943
0.98	0.75539	0.62068	0.61635	0.68337	0.70508	0.47454	0.47471	0.47831	0.47580
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00283	0.00264	0.00344	0.00285	0.00195	0.00195	0.00331	0.00266
0.3		0.00775	0.00698	0.00828	0.00745	0.00489	0.00490	0.00649	0.00574
0.5		0.01630	0.01475	0.01697	0.01545	0.01042	0.01044	0.01218	0.01128
0.7		0.03628	0.03347	0.03642	0.03635	0.02423	0.02427	0.02574	0.02493
0.8		0.06345	0.05983	0.06454	0.06752	0.04537	0.04544	0.04652	0.04584
0.85		0.09128	0.08745	0.09418	0.10049	0.06829	0.06839	0.06894	0.06854
0.9		0.15002	0.14724	0.15697	0.16844	0.12039	0.12051	0.12058	0.12031
0.95		0.31361	0.31694	0.33289	0.35268	0.27215	0.27235	0.27198	0.27201
0.98		0.64345	0.64504	0.68606	0.70573	0.60087	0.60106	0.60201	0.60111
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00033	0.00011	0.00007	0.00065				
0.3		0.00111	0.00041	0.00047	0.00170				
0.5		0.00269	0.00116	0.00137	0.00368				
0.7		0.00599	0.00313	0.00298	0.00870				
0.8		0.01285	0.00886	0.00799	0.01738				
0.85		0.02275	0.01835	0.01756	0.02829				
0.9		0.06116	0.05790	0.05729	0.06669				
0.95		0.23075	0.22637	0.22564	0.24096				
0.98		0.65076	0.63168	0.62752	0.68128				

Table 2a. MSE of the estimators ($T = 100$, $\sigma_{11} = 1$, $\sigma_{22} = 1$, $\sigma_{21} = 0.4$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.002017	0.001243	0.001247	0.001323	0.001297	0.001311	0.001312	0.001387	0.001350
0.3	0.003217	0.002172	0.002189	0.002167	0.002139	0.002221	0.002221	0.002298	0.002261
0.5	0.005631	0.004256	0.004327	0.004115	0.004066	0.004228	0.004228	0.004368	0.004322
0.7	0.013546	0.011940	0.012350	0.011481	0.011374	0.012549	0.012563	0.012704	0.012717
0.8	0.025713	0.025270	0.026281	0.024253	0.024019	0.028315	0.028290	0.028725	0.028694
0.85	0.039454	0.041652	0.043302	0.039902	0.039476	0.051930	0.051917	0.052547	0.052887
0.9	0.068144	0.078019	0.080673	0.074400	0.073074	0.113570	0.113539	0.114974	0.115245
0.95	0.149651	0.189493	0.190981	0.176559	0.171130	0.364919	0.363874	0.366671	0.370330
0.98	0.281292	0.377089	0.371714	0.347022	0.333060	0.955373	0.954624	0.955187	0.965214
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.001236	0.001240	0.001313	0.001286	0.001301	0.001302	0.001372	0.001337
0.3		0.002154	0.002170	0.002149	0.002124	0.002198	0.002197	0.002272	0.002238
0.5		0.004215	0.004280	0.004085	0.004039	0.004168	0.004167	0.004308	0.004259
0.7		0.011821	0.012187	0.011393	0.011294	0.012067	0.012075	0.012299	0.012275
0.8		0.024996	0.025860	0.024074	0.023861	0.026617	0.026591	0.027127	0.027007
0.85		0.041109	0.042443	0.039589	0.039223	0.047266	0.047240	0.048038	0.048198
0.9		0.076936	0.078837	0.073995	0.072768	0.097753	0.097714	0.099600	0.099107
0.95		0.185186	0.184278	0.175467	0.170450	0.276101	0.275634	0.279855	0.280203
0.98		0.366383	0.357841	0.344870	0.331955	0.580129	0.579907	0.588625	0.58581608
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.003196	0.005132	0.002277	0.001219				
0.3		0.005030	0.008322	0.004171	0.002127				
0.5		0.009882	0.015833	0.011597	0.004245				
0.7		0.026314	0.037986	0.031976	0.012995				
0.8		0.053339	0.071840	0.066549	0.027807				
0.85		0.088648	0.120757	0.110548	0.047768				
0.9		0.169981	0.220243	0.219535	0.103255				
0.95		0.368335	0.469592	0.455939	0.242439				
0.98		0.678207	0.846740	0.885145	0.507783				

Table 2b. MSE of the estimators ($T = 100$, $\sigma_{11} = 1$, $\sigma_{22} = 1$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00400	0.00067	0.00066	0.00085	0.00074	0.00070	0.00070	0.00101	0.00085
0.3	0.00629	0.00149	0.00144	0.00151	0.00137	0.00119	0.00119	0.00176	0.00147
0.5	0.01119	0.00363	0.00352	0.00341	0.00317	0.00230	0.00230	0.00325	0.00268
0.7	0.02659	0.01251	0.01261	0.01158	0.01138	0.00739	0.00740	0.00906	0.00817
0.8	0.04816	0.02817	0.02931	0.02643	0.02685	0.01719	0.01725	0.01992	0.01843
0.85	0.07452	0.04963	0.05201	0.04788	0.04911	0.03513	0.03517	0.03852	0.03687
0.9	0.12714	0.09952	0.10456	0.09787	0.10000	0.08603	0.08610	0.09088	0.08939
0.95	0.25791	0.23811	0.24467	0.23708	0.23894	0.31011	0.31034	0.31833	0.31899
0.98	0.48032	0.49519	0.49478	0.49161	0.48799	0.87506	0.87312	0.86814	0.87846
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00064	0.00064	0.00077	0.00069	0.00067	0.00067	0.00091	0.00078
0.3		0.00148	0.00145	0.00150	0.00139	0.00131	0.00131	0.00174	0.00149
0.5		0.00376	0.00370	0.00359	0.00342	0.00301	0.00302	0.00361	0.00321
0.7		0.01305	0.01322	0.01230	0.01211	0.01042	0.01047	0.01127	0.01082
0.8		0.02904	0.03014	0.02744	0.02770	0.02351	0.02362	0.02484	0.02410
0.85		0.05066	0.05278	0.04893	0.04986	0.04327	0.04340	0.04515	0.04416
0.9		0.10046	0.10473	0.09849	0.10024	0.09281	0.09292	0.09582	0.09475
0.95		0.23744	0.24211	0.23662	0.23841	0.26805	0.26803	0.27287	0.27148
0.98		0.49080	0.48830	0.49027	0.48710	0.61926	0.61911	0.62297	0.62315
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00142	0.00235	0.00108	0.00056				
0.3		0.00235	0.00375	0.00217	0.00112				
0.5		0.00490	0.00728	0.00532	0.00253				
0.7		0.01417	0.01786	0.01633	0.01013				
0.8		0.03149	0.03720	0.03762	0.02627				
0.85		0.06016	0.06754	0.06999	0.05609				
0.9		0.12090	0.13212	0.13742	0.11444				
0.95		0.31073	0.34585	0.36571	0.29287				
0.98		0.65212	0.73058	0.76863	0.58960				

Table 2c. MSE of the estimators ($T = 100$, $\sigma_{11} = 1$, $\sigma_{22} = 3$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00073	0.00038	0.00038	0.00042	0.00040	0.00040	0.00040	0.00044	0.00042
0.3	0.00113	0.00065	0.00066	0.00066	0.00065	0.00066	0.00066	0.00070	0.00068
0.5	0.00212	0.00143	0.00146	0.00140	0.00138	0.00140	0.00140	0.00147	0.00144
0.7	0.00485	0.00397	0.00410	0.00377	0.00372	0.00396	0.00396	0.00412	0.00409
0.8	0.00963	0.00867	0.00895	0.00836	0.00832	0.00943	0.00943	0.00963	0.00962
0.85	0.01460	0.01419	0.01472	0.01363	0.01357	0.01667	0.01666	0.01694	0.01702
0.9	0.02467	0.02615	0.02690	0.02510	0.02487	0.03645	0.03650	0.03695	0.03721
0.95	0.05361	0.06453	0.06518	0.06071	0.05921	0.12216	0.12183	0.12244	0.12363
0.98	0.09751	0.12489	0.12452	0.11504	0.11128	0.29552	0.29519	0.29619	0.29824
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00038	0.00038	0.00041	0.00040	0.00040	0.00040	0.00043	0.00041
0.3		0.00064	0.00065	0.00066	0.00065	0.00065	0.00065	0.00069	0.00067
0.5		0.00142	0.00144	0.00139	0.00137	0.00138	0.00138	0.00144	0.00141
0.7		0.00392	0.00404	0.00374	0.00370	0.00381	0.00381	0.00399	0.00394
0.8		0.00858	0.00882	0.00831	0.00828	0.00885	0.00886	0.00912	0.00907
0.85		0.01403	0.01448	0.01355	0.01350	0.01521	0.01521	0.01563	0.01553
0.9		0.02580	0.02632	0.02495	0.02476	0.03133	0.03138	0.03211	0.03196
0.95		0.06308	0.06296	0.06036	0.05898	0.09118	0.09107	0.09200	0.09186
0.98		0.12155	0.11988	0.11446	0.11096	0.18709	0.18707	0.18987	0.18964
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00094	0.00157	0.00070	0.00037				
0.3		0.00152	0.00250	0.00141	0.00064				
0.5		0.00314	0.00508	0.00331	0.00146				
0.7		0.00839	0.01195	0.01005	0.00440				
0.8		0.01930	0.02524	0.02473	0.01109				
0.85		0.02975	0.03822	0.03688	0.01815				
0.9		0.05130	0.06579	0.06489	0.03251				
0.95		0.12558	0.15559	0.16288	0.08494				
0.98		0.21715	0.27185	0.29316	0.15809				

Table 2d. MSE of the estimators ($T = 100$, $\sigma_{11} = 1$, $\sigma_{22} = 3$, $\sigma_{21} = 1.6$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00164	0.00013	0.00013	0.00024	0.00017	0.00015	0.00014	0.00032	0.00023
0.3	0.00261	0.00036	0.00034	0.00041	0.00034	0.00021	0.00021	0.00055	0.00038
0.5	0.00454	0.00105	0.00099	0.00104	0.00096	0.00040	0.00040	0.00102	0.00068
0.7	0.01085	0.00427	0.00431	0.00401	0.00397	0.00133	0.00135	0.00263	0.00190
0.8	0.02002	0.01036	0.01093	0.00978	0.01009	0.00351	0.00353	0.00544	0.00440
0.85	0.02948	0.01784	0.01916	0.01716	0.01788	0.00731	0.00735	0.00979	0.00853
0.9	0.05013	0.03621	0.03904	0.03556	0.03684	0.02068	0.02074	0.02414	0.02257
0.95	0.10207	0.08876	0.09320	0.08913	0.09074	0.09547	0.09538	0.09820	0.09705
0.98	0.18332	0.17809	0.17988	0.17945	0.17969	0.27149	0.27091	0.26891	0.26992
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00012	0.00012	0.00018	0.00014	0.00013	0.00013	0.00024	0.00017
0.3		0.00042	0.00042	0.00044	0.00040	0.00042	0.00042	0.00058	0.00047
0.5		0.00127	0.00126	0.00124	0.00119	0.00116	0.00117	0.00140	0.00124
0.7		0.00484	0.00494	0.00460	0.00456	0.00421	0.00427	0.00458	0.00439
0.8		0.01127	0.01180	0.01069	0.01084	0.00967	0.00977	0.01025	0.00999
0.85		0.01891	0.02005	0.01815	0.01859	0.01643	0.01656	0.01735	0.01700
0.9		0.03733	0.03967	0.03639	0.03732	0.03349	0.03364	0.03497	0.03447
0.95		0.08929	0.09269	0.08927	0.09069	0.09379	0.09389	0.09557	0.09494
0.98		0.17760	0.17849	0.17918	0.17946	0.20580	0.20571	0.20635	0.20607
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00020	0.00031	0.00015	0.00009				
0.3		0.00033	0.00050	0.00030	0.00016				
0.5		0.00062	0.00090	0.00069	0.00038				
0.7		0.00210	0.00250	0.00247	0.00169				
0.8		0.00531	0.00565	0.00599	0.00515				
0.85		0.01023	0.01028	0.01085	0.01073				
0.9		0.02765	0.02710	0.02782	0.02906				
0.95		0.09504	0.09621	0.10006	0.09661				
0.98		0.21216	0.22238	0.23099	0.20384				

Table 2e. MSE of the estimators ($T = 100$, $\sigma_{11} = 3$, $\sigma_{22} = 1$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00676	0.00357	0.00359	0.00383	0.00371	0.00378	0.00379	0.00403	0.00392
0.3	0.01004	0.00598	0.00602	0.00591	0.00580	0.00599	0.00599	0.00627	0.00611
0.5	0.01991	0.01342	0.01356	0.01298	0.01275	0.01299	0.01300	0.01338	0.01327
0.7	0.04433	0.03573	0.03683	0.03430	0.03389	0.03608	0.03605	0.03674	0.03638
0.8	0.08274	0.07491	0.07796	0.07194	0.07155	0.08082	0.08069	0.08214	0.08204
0.85	0.12872	0.12423	0.12896	0.11972	0.11910	0.14432	0.14437	0.14557	0.14640
0.9	0.22864	0.24263	0.25057	0.23384	0.23159	0.32901	0.32924	0.33124	0.33305
0.95	0.49462	0.58939	0.59461	0.56147	0.54882	1.03762	1.03829	1.04945	1.05294
0.98	0.89964	1.15831	1.14806	1.07069	1.03273	2.53653	2.53333	2.53599	2.55101
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00353	0.00355	0.00377	0.00367	0.00373	0.00374	0.00397	0.00386
0.3		0.00591	0.00595	0.00585	0.00575	0.00591	0.00591	0.00618	0.00603
0.5		0.01327	0.01340	0.01287	0.01265	0.01280	0.01281	0.01320	0.01307
0.7		0.03537	0.03632	0.03405	0.03367	0.03465	0.03463	0.03548	0.03508
0.8		0.07410	0.07673	0.07145	0.07112	0.07613	0.07602	0.07780	0.07730
0.85		0.12275	0.12666	0.11889	0.11840	0.13267	0.13267	0.13443	0.13447
0.9		0.23958	0.24566	0.23241	0.23049	0.28680	0.28694	0.29089	0.29054
0.95		0.57797	0.57792	0.55804	0.54655	0.81621	0.81605	0.82558	0.82605
0.98		1.12655	1.10361	1.06429	1.02936	1.67585	1.67548	1.68836	1.68979
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00810	0.01377	0.00632	0.00349				
0.3		0.01418	0.02262	0.01114	0.00579				
0.5		0.02799	0.04516	0.03130	0.01354				
0.7		0.07826	0.10818	0.09717	0.03986				
0.8		0.15632	0.21505	0.20265	0.09026				
0.85		0.26888	0.35645	0.34216	0.15016				
0.9		0.47986	0.61200	0.61873	0.31424				
0.95		1.09031	1.35729	1.38276	0.76379				
0.98		2.25896	2.76499	2.85981	1.63494				

Table 2f. MSE of the estimators ($T = 100$, $\sigma_{11} = 3$, $\sigma_{22} = 1$, $\sigma_{21} = 1.6$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.01477	0.00129	0.00125	0.00176	0.00139	0.00123	0.00121	0.00266	0.00187
0.3	0.02237	0.00337	0.00313	0.00316	0.00277	0.00179	0.00177	0.00422	0.00298
0.5	0.04109	0.01030	0.00973	0.00917	0.00814	0.00386	0.00383	0.00841	0.00582
0.7	0.09574	0.03854	0.03904	0.03388	0.03291	0.01260	0.01260	0.02140	0.01598
0.8	0.17970	0.09381	0.09892	0.08523	0.08709	0.03368	0.03361	0.04516	0.03811
0.85	0.26335	0.15994	0.17209	0.14893	0.15455	0.06670	0.06682	0.08055	0.07251
0.9	0.45339	0.33051	0.35763	0.31722	0.32908	0.19580	0.19619	0.21801	0.20680
0.95	0.92695	0.80420	0.84528	0.80459	0.82004	0.86040	0.86046	0.88267	0.87359
0.98	1.66742	1.61526	1.63489	1.62991	1.63261	2.53650	2.53359	2.52386	2.53538
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00115	0.00115	0.00136	0.00118	0.00114	0.00114	0.00197	0.00140
0.3		0.00373	0.00369	0.00355	0.00339	0.00339	0.00342	0.00453	0.00374
0.5		0.01202	0.01195	0.01127	0.01069	0.01061	0.01070	0.01210	0.01091
0.7		0.04337	0.04439	0.03997	0.03916	0.03855	0.03889	0.04002	0.03879
0.8		0.10180	0.10660	0.09452	0.09520	0.08999	0.09050	0.09108	0.09009
0.85		0.16945	0.18007	0.15907	0.16225	0.15040	0.15108	0.15242	0.15083
0.9		0.34016	0.36241	0.32613	0.33440	0.31390	0.31474	0.32076	0.31717
0.95		0.80834	0.83947	0.80667	0.82003	0.85466	0.85528	0.87038	0.86229
0.98		1.61186	1.62364	1.62758	1.63042	1.88399	1.88385	1.88879	1.88654
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00161	0.00263	0.00123	0.00078				
0.3		0.00288	0.00445	0.00271	0.00136				
0.5		0.00625	0.00894	0.00675	0.00347				
0.7		0.01878	0.02258	0.02264	0.01514				
0.8		0.04927	0.05221	0.05456	0.04609				
0.85		0.09168	0.09154	0.09645	0.09672				
0.9		0.24616	0.24134	0.25380	0.26066				
0.95		0.87493	0.88644	0.92787	0.89750				
0.98		1.93151	2.03331	2.13133	1.90114				

Table 2g. MSE of the estimators ($T = 300$, $\sigma_{11} = 1$, $\sigma_{22} = 1$, $\sigma_{21} = 0.4$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00023	0.00013	0.00013	0.00013	0.00013	0.00013	0.00013	0.00013	0.00013
0.3	0.00038	0.00022	0.00022	0.00022	0.00022	0.00022	0.00022	0.00023	0.00022
0.5	0.00068	0.00044	0.00043	0.00043	0.00043	0.00042	0.00042	0.00043	0.00043
0.7	0.00184	0.00131	0.00131	0.00128	0.00126	0.00122	0.00122	0.00125	0.00124
0.8	0.00398	0.00306	0.00308	0.00297	0.00296	0.00289	0.00289	0.00293	0.00293
0.85	0.00632	0.00519	0.00530	0.00501	0.00501	0.00497	0.00498	0.00507	0.00504
0.9	0.01310	0.01185	0.01220	0.01147	0.01153	0.01213	0.01214	0.01234	0.01241
0.95	0.03743	0.03949	0.04082	0.03784	0.03739	0.04981	0.04977	0.05032	0.05029
0.98	0.11933	0.14467	0.14741	0.13303	0.12838	0.22580	0.22551	0.22738	0.22854
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00013	0.00013	0.00013	0.00013	0.00013	0.00013	0.00013	0.00013
0.3		0.00022	0.00022	0.00022	0.00022	0.00022	0.00022	0.00023	0.00022
0.5		0.00043	0.00043	0.00043	0.00042	0.00042	0.00042	0.00043	0.00043
0.7		0.00130	0.00130	0.00127	0.00125	0.00121	0.00121	0.00124	0.00123
0.8		0.00305	0.00307	0.00297	0.00296	0.00284	0.00284	0.00290	0.00289
0.85		0.00517	0.00526	0.00500	0.00501	0.00483	0.00483	0.00495	0.00491
0.9		0.01178	0.01207	0.01144	0.01151	0.01158	0.01159	0.01181	0.01183
0.95		0.03905	0.04009	0.03773	0.03733	0.04460	0.04457	0.04547	0.04521
0.98		0.14189	0.14313	0.13263	0.12821	0.19246	0.19233	0.19416	0.19438
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00014	0.00018	0.00013	0.00012				
0.3		0.00025	0.00030	0.00025	0.00023				
0.5		0.00049	0.00057	0.00050	0.00045				
0.7		0.00151	0.00167	0.00159	0.00140				
0.8		0.00348	0.00367	0.00369	0.00340				
0.85		0.00599	0.00627	0.00653	0.00574				
0.9		0.01392	0.01462	0.01516	0.01303				
0.95		0.04218	0.04631	0.04877	0.03889				
0.98		0.14351	0.15936	0.16860	0.13080				

Table 2h. MSE of the estimators ($T = 300$, $\sigma_{11} = 1$, $\sigma_{22} = 1$, $\sigma_{21} = 0.8$)

ρ	OLS	FMR with the Kernel Method				FMR with the PW Method			
		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00048	0.00006	0.00006	0.00007	0.00006	0.00006	0.00006	0.00007	0.00006
0.3	0.00077	0.00012	0.00011	0.00012	0.00011	0.00010	0.00010	0.00012	0.00011
0.5	0.00150	0.00029	0.00026	0.00030	0.00027	0.00019	0.00019	0.00024	0.00021
0.7	0.00381	0.00104	0.00094	0.00099	0.00095	0.00057	0.00057	0.00068	0.00061
0.8	0.00788	0.00271	0.00253	0.00263	0.00275	0.00138	0.00138	0.00157	0.00147
0.85	0.01340	0.00558	0.00537	0.00546	0.00593	0.00269	0.00269	0.00310	0.00286
0.9	0.02585	0.01313	0.01302	0.01341	0.01488	0.00659	0.00660	0.00729	0.00698
0.95	0.07116	0.04835	0.05010	0.05099	0.05511	0.03125	0.03125	0.03335	0.03241
0.98	0.20786	0.18046	0.18198	0.19006	0.19461	0.19169	0.19170	0.19403	0.19364
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00006	0.00006	0.00007	0.00006	0.00006	0.00006	0.00007	0.00006
0.3		0.00012	0.00012	0.00013	0.00012	0.00010	0.00010	0.00012	0.00011
0.5		0.00032	0.00029	0.00032	0.00029	0.00023	0.00023	0.00027	0.00025
0.7		0.00115	0.00107	0.00111	0.00108	0.00077	0.00077	0.00086	0.00081
0.8		0.00300	0.00286	0.00293	0.00302	0.00204	0.00204	0.00216	0.00210
0.85		0.00608	0.00593	0.00596	0.00631	0.00421	0.00422	0.00442	0.00430
0.9		0.01405	0.01402	0.01417	0.01535	0.01027	0.01029	0.01063	0.01050
0.95		0.05006	0.05167	0.05182	0.05543	0.04017	0.04021	0.04133	0.04082
0.98		0.18190	0.18321	0.19022	0.19458	0.18373	0.18380	0.18633	0.18553
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00007	0.00008	0.00006	0.00006				
0.3		0.00011	0.00014	0.00011	0.00010				
0.5		0.00023	0.00027	0.00024	0.00022				
0.7		0.00072	0.00078	0.00076	0.00070				
0.8		0.00185	0.00184	0.00189	0.00192				
0.85		0.00372	0.00362	0.00366	0.00399				
0.9		0.00958	0.00891	0.00894	0.01063				
0.95		0.04549	0.04227	0.04193	0.05159				
0.98		0.20109	0.19906	0.20207	0.20564				

Table 2i. MSE of the estimators ($T = 300$, $\sigma_{11} = 1$, $\sigma_{22} = 3$, $\sigma_{21} = 0.8$)

	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00008	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004
0.3	0.00014	0.00007	0.00007	0.00007	0.00007	0.00007	0.00007	0.00007	0.00007
0.5	0.00026	0.00014	0.00014	0.00014	0.00014	0.00013	0.00013	0.00014	0.00014
0.7	0.00070	0.00043	0.00042	0.00042	0.00041	0.00038	0.00038	0.00040	0.00039
0.8	0.00141	0.00097	0.00098	0.00094	0.00094	0.00088	0.00088	0.00091	0.00089
0.85	0.00235	0.00174	0.00176	0.00170	0.00172	0.00161	0.00161	0.00166	0.00165
0.9	0.00473	0.00393	0.00404	0.00382	0.00388	0.00378	0.00378	0.00389	0.00386
0.95	0.01384	0.01346	0.01387	0.01315	0.01318	0.01582	0.01582	0.01618	0.01609
0.98	0.04196	0.04918	0.05026	0.04529	0.04405	0.07280	0.07274	0.07341	0.07367
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00004	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004
0.3		0.00007	0.00007	0.00007	0.00007	0.00007	0.00007	0.00007	0.00007
0.5		0.00014	0.00014	0.00014	0.00014	0.00013	0.00013	0.00014	0.00014
0.7		0.00043	0.00042	0.00042	0.00041	0.00038	0.00038	0.00039	0.00039
0.8		0.00097	0.00097	0.00094	0.00094	0.00086	0.00086	0.00089	0.00088
0.85		0.00174	0.00175	0.00170	0.00173	0.00157	0.00157	0.00162	0.00161
0.9		0.00392	0.00401	0.00381	0.00388	0.00361	0.00361	0.00373	0.00370
0.95		0.01336	0.01368	0.01312	0.01316	0.01425	0.01425	0.01464	0.01451
0.98		0.04825	0.04884	0.04517	0.04400	0.06183	0.06180	0.06240	0.06235
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00005	0.00006	0.00004	0.00004				
0.3		0.00008	0.00010	0.00008	0.00007				
0.5		0.00016	0.00019	0.00016	0.00015				
0.7		0.00047	0.00052	0.00050	0.00046				
0.8		0.00111	0.00119	0.00119	0.00110				
0.85		0.00201	0.00211	0.00218	0.00202				
0.9		0.00461	0.00481	0.00499	0.00446				
0.95		0.01484	0.01564	0.01631	0.01409				
0.98		0.04863	0.05310	0.05622	0.04463				

Table 2j.MSE of the estimators ($T = 300$, $\sigma_{11} = 1$, $\sigma_{22} = 3$, $\sigma_{21} = 1.6$)

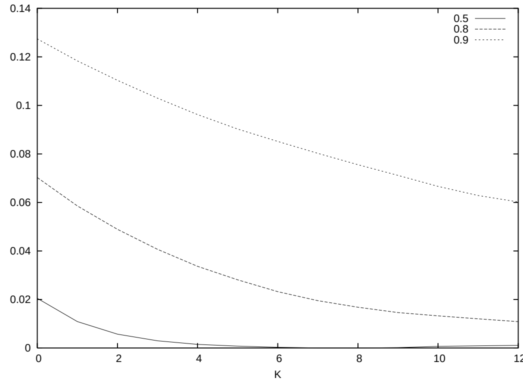
	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00019	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
0.3	0.00031	0.00002	0.00002	0.00003	0.00002	0.00001	0.00001	0.00002	0.00002
0.5	0.00060	0.00007	0.00006	0.00008	0.00006	0.00003	0.00003	0.00005	0.00004
0.7	0.00154	0.00030	0.00025	0.00029	0.00028	0.00008	0.00008	0.00015	0.00011
0.8	0.00316	0.00085	0.00077	0.00082	0.00090	0.00019	0.00019	0.00030	0.00023
0.85	0.00527	0.00176	0.00165	0.00175	0.00200	0.00038	0.00039	0.00058	0.00046
0.9	0.01028	0.00450	0.00449	0.00466	0.00538	0.00100	0.00101	0.00139	0.00118
0.95	0.02875	0.01773	0.01876	0.01909	0.02114	0.00639	0.00640	0.00735	0.00689
0.98	0.08459	0.06826	0.06945	0.07423	0.07704	0.05910	0.05911	0.06055	0.05987
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
0.3		0.00003	0.00003	0.00003	0.00003	0.00002	0.00002	0.00003	0.00003
0.5		0.00009	0.00008	0.00010	0.00009	0.00006	0.00006	0.00008	0.00007
0.7		0.00039	0.00036	0.00038	0.00037	0.00025	0.00026	0.00029	0.00027
0.8		0.00106	0.00101	0.00104	0.00108	0.00072	0.00072	0.00076	0.00074
0.85		0.00212	0.00207	0.00209	0.00226	0.00152	0.00153	0.00157	0.00156
0.9		0.00516	0.00519	0.00520	0.00571	0.00393	0.00395	0.00402	0.00400
0.95		0.01907	0.01996	0.01978	0.02143	0.01556	0.01559	0.01586	0.01578
0.98		0.07014	0.07130	0.07450	0.07709	0.06574	0.06577	0.06682	0.06638
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00001	0.00001	0.00001	0.00001				
0.3		0.00001	0.00002	0.00001	0.00001				
0.5		0.00003	0.00004	0.00003	0.00003				
0.7		0.00010	0.00010	0.00010	0.00010				
0.8		0.00026	0.00026	0.00026	0.00028				
0.85		0.00056	0.00053	0.00053	0.00061				
0.9		0.00167	0.00159	0.00157	0.00183				
0.95		0.01202	0.01153	0.01147	0.01294				
0.98		0.07345	0.07093	0.07047	0.07702				

Table 2k. MSE of the estimators ($T = 300$, $\sigma_{11} = 3$, $\sigma_{22} = 1$, $\sigma_{21} = 0.8$)

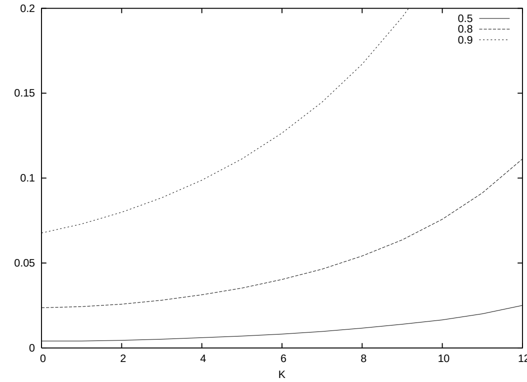
	OLS	FMR with the Kernel Method				FMR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00078	0.00037	0.00037	0.00038	0.00038	0.00037	0.00037	0.00039	0.00038
0.3	0.00119	0.00062	0.00061	0.00062	0.00061	0.00061	0.00061	0.00062	0.00062
0.5	0.00240	0.00129	0.00127	0.00128	0.00125	0.00121	0.00121	0.00125	0.00123
0.7	0.00618	0.00381	0.00378	0.00373	0.00367	0.00347	0.00346	0.00353	0.00353
0.8	0.01293	0.00901	0.00907	0.00875	0.00870	0.00826	0.00826	0.00847	0.00843
0.85	0.02207	0.01636	0.01659	0.01583	0.01597	0.01494	0.01493	0.01529	0.01516
0.9	0.04290	0.03529	0.03610	0.03446	0.03505	0.03472	0.03475	0.03501	0.03501
0.95	0.12910	0.12778	0.13171	0.12376	0.12371	0.15166	0.15177	0.15382	0.15354
0.98	0.40071	0.46383	0.46888	0.43476	0.42291	0.68618	0.68624	0.69317	0.69332
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00037	0.00037	0.00038	0.00038	0.00037	0.00037	0.00039	0.00038
0.3		0.00062	0.00061	0.00062	0.00061	0.00061	0.00061	0.00062	0.00062
0.5		0.00129	0.00127	0.00128	0.00125	0.00121	0.00121	0.00124	0.00123
0.7		0.00381	0.00377	0.00373	0.00368	0.00345	0.00345	0.00352	0.00351
0.8		0.00899	0.00902	0.00875	0.00871	0.00811	0.00812	0.00832	0.00828
0.85		0.01633	0.01650	0.01584	0.01599	0.01462	0.01461	0.01499	0.01485
0.9		0.03516	0.03580	0.03444	0.03504	0.03306	0.03308	0.03354	0.03347
0.95		0.12666	0.12984	0.12345	0.12354	0.13657	0.13662	0.13887	0.13846
0.98		0.45626	0.45709	0.43364	0.42243	0.59365	0.59360	0.60107	0.60046
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00041	0.00052	0.00040	0.00036				
0.3		0.00070	0.00086	0.00070	0.00063				
0.5		0.00143	0.00169	0.00147	0.00133				
0.7		0.00429	0.00475	0.00466	0.00410				
0.8		0.01036	0.01094	0.01131	0.01017				
0.85		0.01901	0.01941	0.01992	0.01869				
0.9		0.04082	0.04196	0.04366	0.04028				
0.95		0.14264	0.15103	0.15715	0.13334				
0.98		0.47310	0.51733	0.55325	0.43255				

Table 2l. MSE of the estimators ($T = 300$, $\sigma_{11} = 3$, $\sigma_{22} = 1$, $\sigma_{21} = 1.6$)

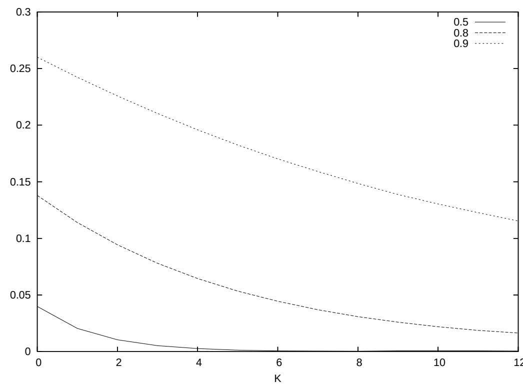
ρ	OLS	FMR with the Kernel Method				FMR with the PW Method			
		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1	0.00174	0.00009	0.00008	0.00011	0.00009	0.00008	0.00008	0.00013	0.00010
0.3	0.00277	0.00021	0.00018	0.00022	0.00019	0.00012	0.00012	0.00022	0.00017
0.5	0.00537	0.00064	0.00053	0.00065	0.00052	0.00024	0.00024	0.00047	0.00033
0.7	0.01391	0.00264	0.00221	0.00246	0.00231	0.00070	0.00070	0.00115	0.00085
0.8	0.02923	0.00825	0.00746	0.00770	0.00813	0.00196	0.00194	0.00325	0.00237
0.85	0.04696	0.01537	0.01441	0.01496	0.01691	0.00348	0.00347	0.00461	0.00384
0.9	0.09287	0.04124	0.04131	0.04156	0.04771	0.00958	0.00956	0.01211	0.01069
0.95	0.26225	0.16251	0.17266	0.17290	0.19183	0.05916	0.05917	0.06598	0.06279
0.98	0.76244	0.61612	0.62670	0.66714	0.69265	0.54222	0.54230	0.55300	0.54853
		CCR with the Kernel Method				CCR with the PW Method			
ρ		BA(AN)	QS(AN)	BA(NW)	QS(NW)	BA(AN)	QS(AN)	BA(NW)	QS(NW)
0.1		0.00009	0.00008	0.00010	0.00009	0.00008	0.00008	0.00012	0.00009
0.3		0.00026	0.00024	0.00027	0.00024	0.00019	0.00019	0.00027	0.00022
0.5		0.00085	0.00077	0.00086	0.00076	0.00057	0.00057	0.00071	0.00062
0.7		0.00345	0.00318	0.00332	0.00319	0.00226	0.00226	0.00246	0.00232
0.8		0.01014	0.00966	0.00972	0.00998	0.00719	0.00721	0.00748	0.00728
0.85		0.01858	0.01807	0.01814	0.01941	0.01341	0.01345	0.01355	0.01347
0.9		0.04707	0.04755	0.04673	0.05097	0.03657	0.03665	0.03675	0.03665
0.95		0.17468	0.18350	0.17956	0.19461	0.14412	0.14432	0.14515	0.14487
0.98		0.63312	0.64374	0.66981	0.69322	0.59577	0.59603	0.60281	0.59988
		DOLS							
ρ		GS(01)	GS(05)	AIC	BIC				
0.1		0.00008	0.00010	0.00008	0.00007				
0.3		0.00013	0.00016	0.00014	0.00013				
0.5		0.00028	0.00033	0.00030	0.00027				
0.7		0.00090	0.00093	0.00095	0.00090				
0.8		0.00257	0.00253	0.00255	0.00273				
0.85		0.00503	0.00482	0.00485	0.00544				
0.9		0.01511	0.01430	0.01418	0.01649				
0.95		0.11078	0.10646	0.10573	0.11957				
0.98		0.65965	0.63656	0.63324	0.69200				



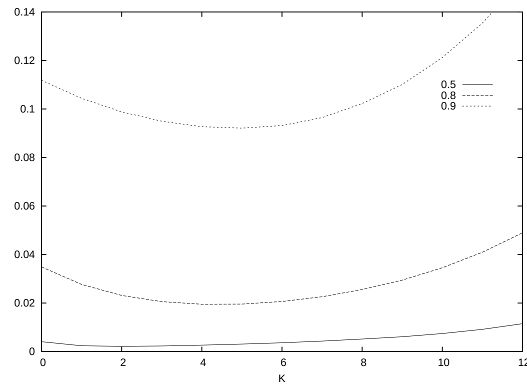
(a) The bias ($T = 100$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{21} = 0.4$)



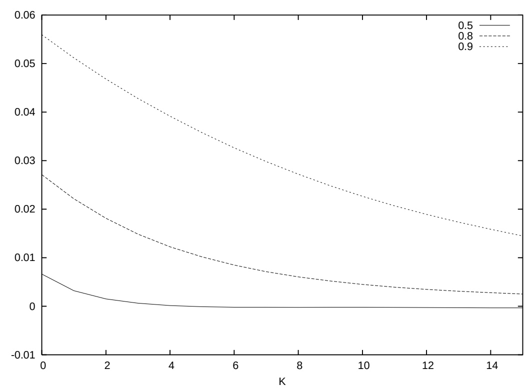
(b) The MSE ($T = 100$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{21} = 0.4$)



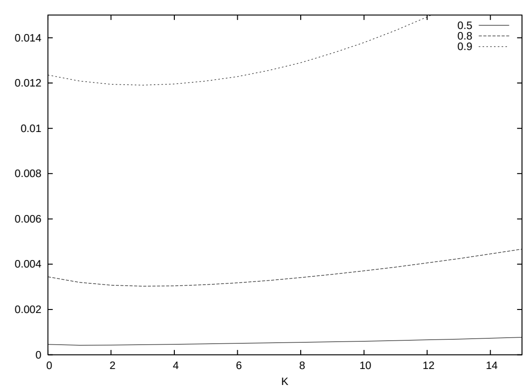
(c) The bias ($T = 100$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{21} = 0.8$)



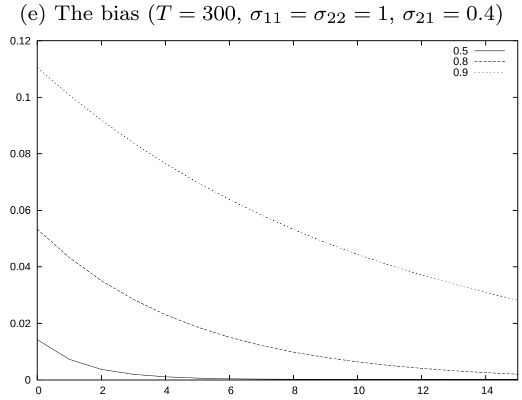
(d) The MSE ($T = 100$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{21} = 0.8$)



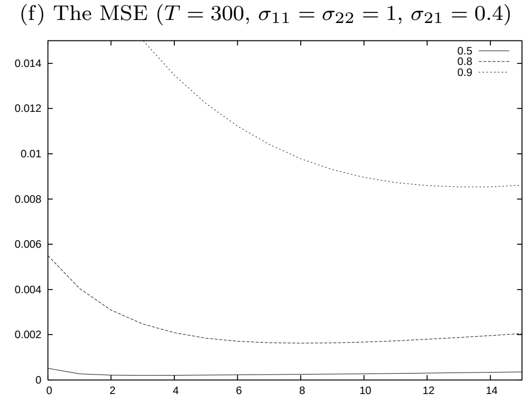
(e) The bias ($T = 300$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{21} = 0.4$)



(f) The MSE ($T = 300$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{21} = 0.4$)

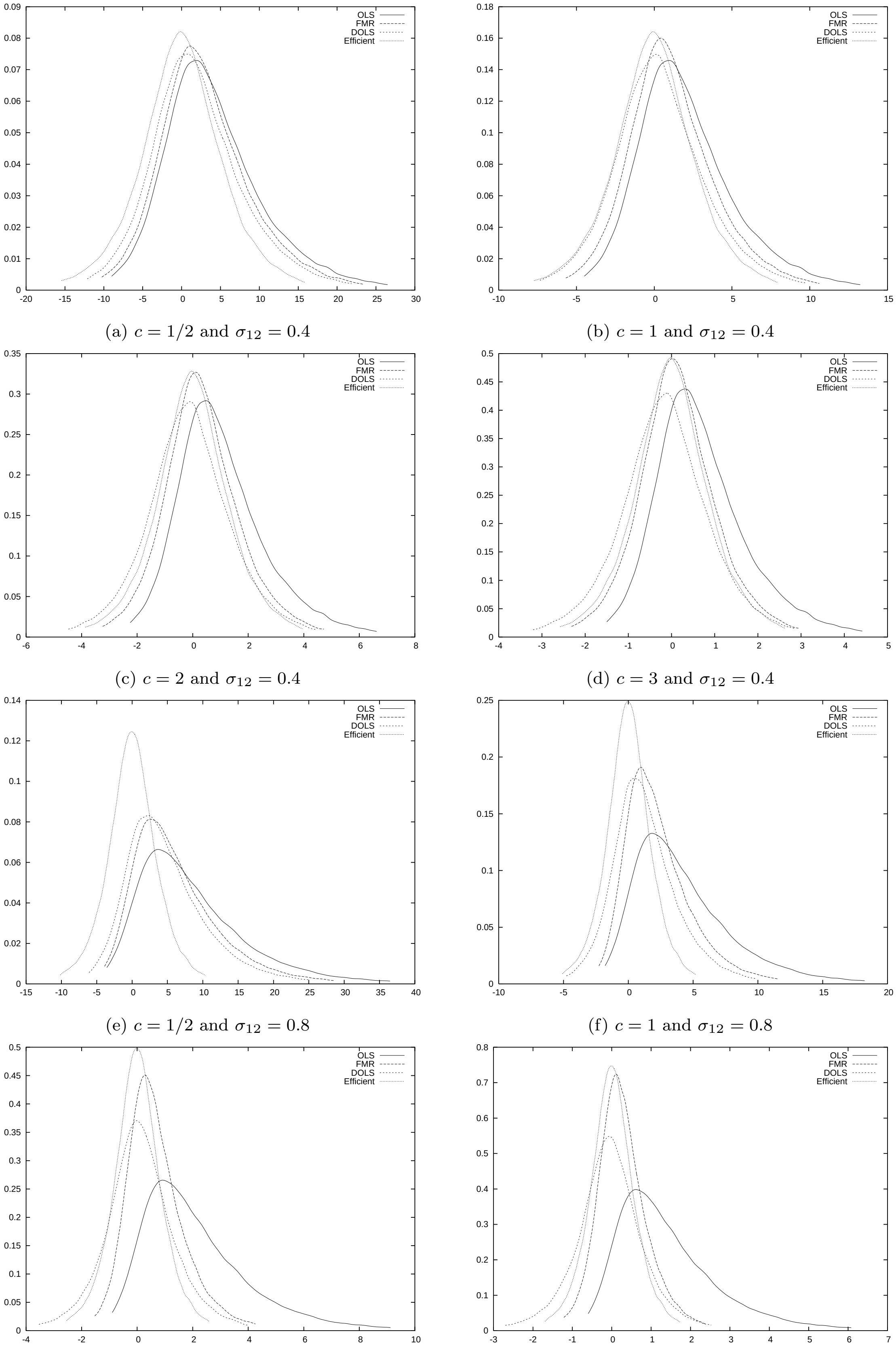


(g) The bias ($T = 300$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{21} = 0.8$)



(h) The MSE ($T = 300$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{21} = 0.8$)

Figure 1: The finite sample bias and MSE of the DOLS estimator



(a) $c = 1/2$ and $\sigma_{12} = 0.4$

(b) $c = 1$ and $\sigma_{12} = 0.4$

(c) $c = 2$ and $\sigma_{12} = 0.4$

(d) $c = 3$ and $\sigma_{12} = 0.4$

(e) $c = 1/2$ and $\sigma_{12} = 0.8$

(f) $c = 1$ and $\sigma_{12} = 0.8$

(g) $c = 2$ and $\sigma_{12} = 0.8$

(h) $c = 3$ and $\sigma_{12} = 0.8$

Figure 2: Probability density functions of the estimators